

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2002, 18:27:05 ; Search time 4750.6 Seconds  
(without alignments)  
9613.790 Million cell updates/sec

Title: US-09-523-647-1

Perfect score: 2111

Sequence: 1 ttgtcaggtagtagtagaga.....catcaaatatcatcccgta 2111

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main.\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*
- 23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq.\*
- 25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq.\*
- 26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq.\*
- 27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq.\*
- 28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq.\*
- 29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq.\*
- 30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq.\*
- 31: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq.\*
- 32: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq.\*
- 33: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq.\*
- 34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq.\*
- 35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq.\*
- 36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq.\*
- 37: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq.\*
- 38: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq.\*
- 39: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq.\*
- 40: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq.\*
- 41: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq.\*
- 42: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq.\*
- 43: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq.\*

- 44: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq.\*
- 45: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq.\*
- 46: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq.\*
- 47: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq.\*
- 48: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq.\*
- 49: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq.\*
- 50: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq.\*
- 51: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq.\*
- 52: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq.\*
- 53: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq.\*
- 54: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq.\*
- 55: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq.\*
- 56: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq.\*
- 57: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq.\*
- 58: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq.\*
- 59: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq.\*
- 60: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq.\*
- 61: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq.\*
- 62: /cgn2\_6/ptodata/1/pna/US6023\_COMB.seq.\*
- 63: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq.\*
- 64: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq.\*
- 65: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq.\*
- 66: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq.\*
- 67: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq.\*
- 68: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq.\*
- 69: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq.\*
- 70: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq.\*
- 71: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq.\*
- 72: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq.\*
- 73: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq.\*
- 74: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq.\*
- 75: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	IDB	ID	Description
1	2111	100.0	2111	19	US-09-523-647-1	Sequence 1, Appli
C 2	2106.2	99.8	1230230	18	US-09-438-185-1	Sequence 1, Appli
C 3	2106.2	99.8	1230230	18	US-09-438-185A-1	Sequence 378, Appli
4	1671	79.2	1671	32	US-09-841-132-378	Sequence 417, App
5	868	41.1	1659	32	US-09-841-132-417	Sequence 131, Appl
C 6	767.6	36.4	1947	37	US-10-007-693-131	Sequence 132, Appl
C 7	719.6	34.1	1278	37	US-10-007-693-132	Sequence 21, Appli
8	363.6	17.2	731	37	US-10-007-693-21	Sequence 63, Appl
9	179.4	8.5	269	18	US-09-410-568-63	Sequence 63, Appl
10	179.4	8.5	269	18	US-09-426-571-63	Sequence 63, Appl
11	179.4	8.5	269	18	US-09-454-684-63	Sequence 63, Appl
12	179.4	8.5	269	18	US-09-454-684A-63	Sequence 63, Appl
13	179.4	8.5	269	22	US-09-556-877-63	Sequence 63, Appl
14	179.4	8.5	269	22	US-09-598-419-63	Sequence 63, Appl
15	179.4	8.5	269	24	US-09-620-412A-63	Sequence 63, Appl
16	179.4	8.5	269	32	US-09-841-132-63	Sequence 63, Appl
17	120.8	5.7	800	33	US-09-868-987-2	Sequence 2, Appli
18	120.8	5.7	800	33	US-09-868-987-2	Sequence 2, Appli
19	48.4	2.3	414	26	US-09-666-355A-738	Sequence 738, Appl
C 20	48	2.3	6033	1	PCT-US99-26796-285	Sequence 285, App
C 21	46.8	2.2	391	23	US-09-619-643-12199	Sequence 12199, A
C 22	46.8	2.2	391	25	US-09-654-617-340068	Sequence 340068,
C 23	46.8	2.2	391	27	US-09-684-016-340068	Sequence 340068,
C 24	46.8	2.2	391	53	US-60-145-485-6512	Sequence 6512, Ap
C 25	46	2.2	48	18	US-09-461-705-20	Sequence 20, Appl
C 26	45.6	2.2	7218	8	US-08-466-194-14	Sequence 14, Appl
C 27	45	2.1	2304	26	US-09-662-254A-43	Sequence 43, Appl
C 28	45	2.1	50000	26	US-09-662-254-24	Sequence 24, Appl
C 29	45	2.1	50000	26	US-09-662-254A-24	Sequence 24, Appl
C 30	44.8	2.1	543029	19	US-09-528-237A-1549	Sequence 1549, Ap
31	44.6	2.1	2000	33	US-09-887-272A-5263	Sequence 5263, Ap





```
QY 1021 agaagaggccaataactaactaactgctgctactgtaacttactgctggtggacacaaatgtct 1080
Db 641979 AGAAGAGGTCAAAATCACTAACCTTCTACTGTAACCTTACTCGGTGGACACAAATGTGTCT 641920
QY 1081 gcaaatgaactacacagttgtaataagccttgggtacaaagtaaatatctctggtgctgcat 1140
Db 641919 GCAAAATGAATACAGTCTGTAATCAGCCTTGTGTACAAGTAATAATATCTCTGGTGTCTGAT 641860
QY 1141 tggcttactgatatgaacctgtgagtaactctatctcgaatctcgaactgagacttg 1200
Db 641859 TGGCTTACGTATGTAACCTGTGGAGTACTCTATCTCGATCGAATCCCGGAGACTTG 641800
QY 1201 gtcttcctgatgtgctgacaaagatacactccctctctggtgttacagttactgaagct 1260
Db 641799 GTTCTTTCATGTGCTGATCAACAGATACACTCCCTCTCTGCTGTGTACAGTACTCGAAGCT 641740
QY 1261 cctgggtggagagatcgtgtaataaagctgttggcgctattaaagaaatgtgcccaggga 1320
Db 641739 CCTGGTGGAGAGATCTGCTGTATAAAGCTTGTGTGGCGTATTAAAGAAATGTGCCAGGA 641680
QY 1321 gaacccctccagtttaaaactttagtgaagctcaagttcctcgaagattcacaatacaa 1380
Db 641679 GAACCCCTCCAGTTTAAACTTGTAGTGAAGCTCAAGTCTCTGGAAGATTCACAAATCAA 641620
QY 1381 gtgcagtaactagtagtctaacctgcggaacatgtacatctctgcgagaaacacaaaca 1440
Db 641619 GTTGCAGTAACTAGTAGTGAAGTCAACTGCGGAACATGTACATCTTGGCAGAAACAAACAA 641560
QY 1441 cattggaagggtctgcagctaccatcatgtgctgattagacacaaatgatcctctctgt 1500
Db 641559 CATTTGGAAGAGTCTTGCAGCTACCCATATGTGCGTATTAGACACAAATGATCTCTATCTGT 641500
QY 1501 gtaggagaaataactgtctatcgtatcgtgtgaactaacctggtctctcgtcgaagatact 1560
Db 641499 GTAGGAGAAATACTGTCTATCTGATCTGTAACCTGATACCTGCTTCTGCTGAAGATACT 641440
QY 1561 aagctatcttbaacttgaagttctcaaaagaacttcaagcaaatagctctctcaggtccca 1620
Db 641439 AAGCTATCTTAACTCTGAAGTCTCAAAAGAATCTCAAGCCAAATAGCTCTCTCAAGTCCA 641380
QY 1621 actaaaggaaacgatttcaggttaataccgtgttcttcgacgctttacctaactcgggtct 1680
Db 641379 ACTAAAGGAACGATTTCAAGTAATACCGTGTGTTTTCGACGCTTTACCTTAAACTCGGTCT 641320
QY 1681 aaggaatctgtagagtttctgttaccctgaaaggatgtctcccgagagatgctcgcggc 1740
Db 641319 AAGGAATCTGTAGAGTCTTCTTACTTACCTTGAAAGGATTTGCTCCGGAGATGCTCGCGC 641260
QY 1741 gaagctattcttctctgatacacactgacttccaccagtatcagacacagaaataaccac 1800
Db 641259 GAAGCTATTCTTCTTCTGTATACACTGACTTCAACAGTATACCCAGTATCAGACACAGAAATACCCAC 641200
QY 1801 gtgtataaattctaaaggaaattactcctaaagcagagatattccgctctgctttaggat 1860
Db 641199 GTGTATTAATTTCTAAGGAATTTATCTTAAAGCAGACGATATTCGCTCTGCTTTTAGGAT 641140
QY 1861 agctttcaagaagtaaccgctttagtaccttacgtactaaagcggtttttttgttttata 1920
Db 641139 AGCTTTCAAGAAGTACCCCTTTAGTACCTTACGTACTTAAGCGGTTTTTTTGTATTATA 641080
QY 1921 agctttcaatccaatcgttagagtttcttataatcaagaatatattaaagtttctgaaatc 1980
Db 641079 AGCTTTCAATCCCAATCGTAGAGTCTTCTAATCAAGATATATTAAAGTTTCTGAAATC 641020
QY 1981 ctaagatttatatttaaagcccatcttttttaggtatgtaataataaatttttttaataagct 2040
Db 641019 CTAAAGATTTATTTAAAGGCCCATCTTTTGTAGGTATGTAATTAATAATTTTAAATTAATGT 640960
QY 2041 ttctcctagtgtaacctgctctcttaggaactacactaggaacagcggtatgctacaaatc 2100
Db 640959 TTTCTTAGTGTAACTGCTTCTTTAGGAACTACACTAGGAGAACGGTATGTATCAATAATC 640900
```

```
QY 2101 tacatcccgta 2111
Db 640899 TACATCCCGTA 640889

RESULT 3
US-09-438-185A-1/c
; Sequence 1, Application US/09438185A
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1
```

```
Query Match 99.8%; Score 2106.2; DB 18; Length 1230230;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2108; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttgatccagtgattagagatgaatttaattcctgaacttaacttaattcaagataaataaacc 60
Db 642999 TTGATCGGGTAGTTAGGAGATGAATTAATTCCTTGACTTACCTTTAATTCAGATATAAAT 642940
QY 61 aaatgttgagggttaagagttttacaaaacattctaccacgatggcagagaagaaaaataaac 120
Db 642939 AAATGTTGAGGGTAAAGAGTTTACAAAACATTTACCCGATGGCAGAGAAAAAATAAAT 642880
QY 121 atgcgataggagatcccttatgtccaaactcatcagacgagtagttacggtcccttgcgcta 180
Db 642879 ATGCGCATAGGAGATCCCTATCTCCAAACTCATATCAGACGAGTACTTACGCTCTCCGCTA 642820
QY 181 acgaatatacgaagttacttttccagcgggggatagagccgctgtacagagctctctg 240
Db 642819 ACGAGTATGGCGAGTTGCTTTGCCGCGGGGTATAGAGCCCGCTGTAGCAGAGTCTCTG 642760
QY 241 attactaagatcgtctcgttagtgcggaacaaagcccgacactgttctctatgcagcgaag 300
Db 642759 ATTACTAAGATCGTCGTAGTGGGAAACAAAGCCAGCACCTCTTCTTATGACAGCGAAG 642700
QY 301 aaggttagacttgccttagaataaaacacccagttgaacaaaagccggtggtctttt 360
Db 642699 AAGGTTTACACTTGTCCGTAGAAATAAACACCACTTTGAACAAAAAGCCCTGCTGCTTTT 642640
QY 361 tctaataaagaatttttccctgtgaagagagacgatgcaacctgaagctcagcaa 420
Db 642639 TGTGATAAAGAAATTTTATCCCTGTGAAGAGGAGCATGTCAACCTGTAGAGGCTACGAA 642580
QY 421 gagcttctcagdaagattgtattcttctgtataaagtaaacgatgattgcaacgttagaatt 480
Db 642579 GAGCTTCTCTCGGAAGAATTTGTATTCTGTAAAAGTAACAGATGATTTGCAACCGCTAGAAAT 642520
QY 481 tgcagtcgcttccagataacgctactgttaggatcctcttaacctattgaaaaccttgcct 540
Db 642519 TGCCAGTCCGTTCCAGAAATACGCTACTGTAGGATCTCTTACCTTATACCTTATTCATTCCT 642460
QY 541 ataggcaaaaagattgttggatgttggattacacaacacagctacacttgcgaagctgaa 600
```



642459	ATAGGCAAAAAAGATGTGTGTGATTTACACACAGCTACCTTGGCAAGCTGAA	642400
601	ttcgttaacgactgataccgagaacaaacctcctacaagtgtatgggaaattagctctgaaaatc	560
642399	TTTCGTAAAGCAGTGATCCAGAAACAACTCCTACAAAGTGATGGGAAATTAGTCTGGAAATC	642340
661	gataccctgggtgcaggagataaaatgcataaattactgtatgggtataaacctcttaagaa	720
642339	GATCGCCCTGGGTGCAGGAGATTAATGCAAAATTACTGTATGGGTAAAAACCTCTTAAAGAA	642280
721	ggttgctgtctcacagctgctactgtatgtctgtgccagagctccgttttactactaa	780
642279	GGTGTGCTTCCACAGCTGCTACTGTATGTGCTTTGCCAGAGCTCCGCTTCTTATACTTAA	642220
781	tgccgttcaaccagccattgtattaaagcaagaagacctgactgtctgtgctctaagatgc	840
642219	TGCGGTCAACACGCCATTGTATTAAAGCAAGAGGACCTGACTGTGCTTGCTTAAGATGC	642160
841	ccgtatgtctacaaaatcgaagttagaacacagatctgtctattgtccgttaacgttaact	900
642159	CCGTATGTCTACAAAATCGAAAGTAGTGAACACAGATCTGTATTGCCCCGTAAAGCTAACT	642100
901	gtagataatcctgttcccatggctattctcatgtcatctgtgtcctaaagagttctctctttt	960
642099	GTAGATAATCCTGTTCCCGATGGCTATTCTCATGCATCTGGTCAAGAGAGTTCTCTCTTTT	642040
961	aacttagagacatgagacctggcgataaaaggattttacaagttgaattctgcacctcaa	1020
642039	AACTTAGGAGACATGAGACCTGGCGATAAAAGGATATTACAGTTGAGTTCTGCCCTCAA	641980
1021	agaagaggtcaaatcactaaacgtgtctactgttaactactgctgcgtgggacacaaatgtct	1080
641979	AGAAGAGGTCAAAATCACTAACGTTGCTACTGTAACTTACTGCGGTGGACACAAATGTCT	641920
1081	gcaaatgtaactacagttgttaatgagcctgtgtacaaagtaaatctcgtgtgctgat	1140
641919	GCAAAATGTAACACAGTTGTTAATGAGCCTGTGTGTACAAAGTAATATCTCTGCTGCTGAT	641860
1141	tggtctacgtatataaacctgtggagtaactctatctcagtatcgaatcctggagacttg	1200
641859	TGGTCTTAGTATGTATGAACCTGTGGAGTACTCTATCTCAGTATCGAATCTGGGAGACTTG	641800
1201	gtcttcattgctgtcgtgatcccaagatacaactccctctgggttgcagtaactcgaagct	1260
641799	GTTCCTCATGATGTCTGTATCCAAAGATACATCCCTTCTTGGTGTTCAGTACTCGAAGCT	641740
1261	cctggtggagagatctgctgtaataaagtgtgttgcgataaagaataatgtgcccaagga	1320
641739	CTGTGTGGAGAGATCTGCTGTATTAAGTTGTTTGGCGTATTAAAGAAATGTGCCCAGGA	641680
1321	gaaacccctccagtttaaactctgtgaaagctcaagttcctggaagattcaacaatcaa	1380
641679	GAACCCCTCCAGTTTAAACTGTAGTGAAGCTCAAGTTCCTGGGAAGATTCCAAANTCAA	641620
1381	gttgacgataactagtcactgaactgcggacaatgtacatctgtcgcgagaacaacaaca	1440
641619	GTTCGAGTAATAGTGAAGTCTAACTGCGGAACATGTATACATCTTGGCGAGAAACAACA	641560
1441	cattggaaaggtcttcagctaccccatatgtgcgtatttagacacaaatgactctatcgt	1500
641559	CAITTGGAAAGGTCTTGCAGAGTACCCCATATGTGGGTATTAGACACAAATGATCCTATCTGT	641500
1501	gtaggagaaaatactgctatcgtatctgtgtaactaacccgtggttctgctgaagatact	1560
641499	GTAGGAGAAAATACGTGCTATCGTATCTGTGTAACCTAACCGTGTCTGCTGAAGATACT	641440
1561	aacgtatctttaacttgaagttctcaaaaagaacttcagccaatagctcttcaggtcca	1620
641439	AACTGATCTTAACTTTGAAGTTCTCAAAAAGAACTTCAGCCAATAGCTTCTTCAGGTCCA	641380
1621	actaaagaaacgatttcaggtaataaccgtgttttcgaagctttactaaactcggttct	1680
641379	ACTAAAGAAAGCATTTACGTAATACCGTGTGTTTCACGCGTTTACCTAAACTCGTGTCT	641320

QY	1681	aaggaaatctgtagaagtttcttggttacacctttaaagtattgctcccgagagatgctcgcgc	1740
DG	641319	AAGGAATCTGTAGAGTTTCTTGTTACCTTGAAGGTATTGGTCGCCGAGATGCTCGCGC	641260
QY	1741	gaagctattcttcttcctgtatcacactgacttaccagttacagacagaaaatacccac	1800
DG	641259	GAAGCTATTCTTCTTGATACACTGACTTACCACGATTCAGACACAGAAATAACCCAC	641200
QY	1801	gtatattaattctaaggaattatctaaagcagcgcgatattccgtctgcttaggat	1860
DG	641199	GTGTATTAAATCTTAGGAATTATCTTAAGCAGAGCGATATTCGCCTCTCTTAGGAT	641140
QY	1861	agcttccaagaagtagccgccttagtacottactactaaagcgggttttttgtttata	1920
DG	641139	AGCTTTCAAAAGAAGTACCOCCTTAGTACCTTACGTACTTAAGCGGTTTTTGTGTTATA	641080
QY	1921	agctctcaatccaatcgtagagttcttctaacaagatatatttaagttttcgaaatc	1980
DG	641079	AGCTTCTCAATCCAATCGTAGAGTTCTTAATCAAAGATATTATTTAAGTTTCTGAAAATC	641020
QY	1981	ctaagattattttaaacgcccatcttttagtgatgtaattaaatttttaattaaagt	2040
DG	641019	CTAAGATTATTPTTAAAGGCCATCTTTTGTGATGTAATTTAAATTTTAATTAATGT	640960
QY	2041	tttctagtgtaacctgcttcttttaggaactacactaggagaacggtatgctcaaatc	2100
DG	640959	TFTCTAGTGAACCTGCTCTTTTAGGAACCTAGACTAGGAGAACGGTATGTCATCAATC	640900
QY	2101	tacatccccgta 2111 	
DG	640899	TACATCCCCGTA 640889	
RESULT 4			
US-09-841-132-378			
; Sequence 378, Application US/09841132			
; GENERAL INFORMATION:			
; APPLICANT: Bhatia, Ajay			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Probst, Peter			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND			
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION			
; FILE REFERENCE: 210121.469C8			
; CURRENT APPLICATION NUMBER: US/09/841.132			
; CURRENT FILING DATE: 2001-04-23			
; NUMBER OF SEQ ID NOS: 599			
; SOFTWARE: FastSeq for Windows Version 3.0/4.0			
; SEQ ID NO 378			
; LENGTH: 1671			
; TYPE: DNA			
; ORGANISM: Chlamydia pneumoniae			
US-09-841-132-378			
Query Match 79.2%; Score 1671; DB 32; Length 1671;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1671; Conservative 0; Mismatches 0; Indels 0; Gaps			
QY	139	atgtccaactcatcagacgagtagttacggtctctgcttaacgagtagtgcgagttgc	198
DG	1	atgtccaactcatcagacgagtagttacggtctctgcttaacgagtagtgcgagttgc	60
QY	199	tttgcacgcggttatagagccgtctgtagcaggtctctgattactaagatcgctcgt	258
DG	61	tttgcacgcggttatagagccgtctgtagcaggtctctgattactaagatcgctcgt	120
QY	259	atgtcggaacaaacccagcagcctgttctctatgacagcgaaggttagactgtccgt	318
DG	121	atgtcggaacaaacccagcagcctgttctctatgacagcgaaggttagactgtccgt	180
QY	319	agaaataaacaccagtttgaaacaaaaaacggtgtgttttgtgtgataagaatttat	378

Query Match	79.2%;	Score 1671;	DB 32;	Length 1671;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1671:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]





```
QY 943 caaagattctcttcttaacttagagacatgagacctggcgataaaaaaggtattaca 1002
Db 1527 CACCGTGACTGACGTTTACTCTTGGAGATATGCAACCTGGAGACGACACAACTACT 1468
QY 1003 gttgagttctgcccctcaaaagaaggtcacaactcaactaactgctgtaacttaetgc 1062
Db 1467 GTAGAGTTTGTGCGCTTAAACGGTGGTGGCTACCAATATAGCAACGGTTTCTACTGT 1408
QY 1063 gdtgacacaaatgtctgcaaatgtaactacagttgtttaaagccctgtgtgtaacaagta 1122
Db 1407 GGAGGACATAAAATACAGCAAGCGTAACAACCTGTGATCAATGAGCCCTTGGCTACAAGTA 1348
QY 1123 aatatctctggtgctgattggtctacgtatgtaaacctgtgagtagtctatctcagta 1182
Db 1347 ACTATTGCGAGGAGCAGATTGCTCTATGTTTGTAAAGCCGTAGATAATGTGATCTCCGTT 1288
QY 1183 tcsaatctgtgagacttggtcttctcatgattctggtatccaagatacaactccctctggt 1242
Db 1287 TCCATCTCGGAGATCTTGTGTCGAGATGTCGCTGTTGAAGACACTCTTCTCCCGGA 1228
QY 1243 gtacagtaactcgaagctcctggtgagagatctgctgtaataaagttggttgcgtatt 1302
Db 1227 GTCACAGTCTTGAAGCTGCGAGGAGCTCAAAATTTCTTGTAAATAAGTAGTTGGACGTG 1168
QY 1303 aaagaaatgtcccagagagaacccctccagtttaaaacttgtgtgaaagctcaagttcct 1362
Db 1167 AAAGAACTCAATCTCGGAGAGCTCTACAGTATAAAGTTCTAGTAAGAGCACAACTCCT 1108
QY 1363 ggaagattcacaaatcaagttgcagtaactagtgagcttaactgcggaacatgtacatct 1422
Db 1107 GGACAATTCACAAATTAATGTTGTTGTAAGAGCTGCTCTGACTGTGGTACTTGTACTTCT 1048
QY 1423 tgcgcagaaacaaacacatggaaggtctgtgagcttgcagctaccocatgtgctattagac 1482
Db 1047 TCGCGAGAGCGCAACACTTACTGGAAGGAGTGTGCTACTCATATGTGCGTAGTAGAT 988
QY 1483 acaaatgatctatctgtgtaggaataactctctatctatctgctgtgtaactaaacct 1542
Db 987 ACTTGTGACCCCTGTTGTGTAGGAGAAATACGTTTACCGTATTGTGTGTCACCAACAGA 928
QY 1543 gttctgtcgaagataactaactatctttaaacttgaagttctcctcaaaagaacttcagcca 1602
Db 927 GGTCTGCGAAGATACAAATGTTCTTAAATGCTTAAATCTCTAAAGAACTGCAACCT 868
QY 1603 atagttctcaggttccaactaaaggaaacgatttcagggttaacacctgttttcgacgct 1662
Db 867 GTATCCTTCTCTGGACCAACTTAAAGGAACGATTACAGGCAATACAGTAGTATTCGATTCTG 808
QY 1663 ttacctaaactcgggtcttaaggaatctgttagagtttctgttaccttgaagggtattgct 1722
Db 807 TTACCTAGATTAGGTTCTTAAAGAACTGTAGAGTCTTCTGTAACTTGAAGCAGTATCA 748
QY 1723 ccgagagatgctcggcggaagctattcttcttctgtataactgaactgaactcaccagttaca 1782
Db 747 GCTGGAGATGCTGCTGGGAAGCGATCTTCTTCGATACATGATGCTTCCAGTTCT 688
QY 1783 gacacagaaataaccacgtgtattaaattctaa 1816
Db 687 GATACAGAAATACACATCTATTAAATCTTTGA 654
```

## RESULT 7

```
; US-10-007-693-132/C
; Sequence 132, Application US/10007693
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
```

```
; SEQ ID NO 132
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-10-007-693-132
```

```
Query Match 34.1%; Score 719.6; DB 37; Length 1278;
Best Local Similarity 74.3%; Pred. No. 2.1e-181;
Matches 908; Conservative 0; Mismatches 314; Indels 0; Gaps 0;
```

```
QY 595 gctgaattctgaagcagtgatccagaaacaactcctcaagtgatgaggaattagctcgg 654
Db 1278 GCAGAGTTCGTAGCGAGTGATCCAGCGACAACCTCTACTGCTGATGCTAAGCTAGTTGG 1219
QY 655 aaaaatcgatcgctgggtcgaggataataatgaataaacttactgtatgggtaaaaactctt 714
Db 1218 AAAATTCACCGCTTAGACAAAGCGAAGAGTAAATTTACTGTATGCGGTAAANCCCTCT 1159
QY 715 aaagaaggttgcgtctccagcgtcgtactgtatgtctgtcccgagctccgtctctat 774
Db 1158 AAAGAAGGTTGCTGCTTTACAGCTGCAACAGTATGCGCTTGTCCAGAGATCCCTTCGGTT 1099
QY 775 actaaatgcggtcaaccgccatttgttaagcaagaagacacctgactgctgtcctca 834
Db 1098 ACAAAATGTGGACAACCTGCTATCTGTGTTAAACAAGAGCCACAGAAATGCTGTTTG 1039
QY 835 agatgcctctgtatgtacaaaatcgaaagtagtgacacagagatctgtctattgcccgaac 894
Db 1038 CGTTGCCCACTAGCTTTACAAAATTAATGTAGTGAACCAAGGACAGCAATAGCTCGTAAC 979
QY 895 gtaactatagataactcctgtcccgatgctatctcatcgtcgtcctcaagagttctc 954
Db 978 GTTGTGTTGAAATCCTGTTCCAGATGGTTACGCTCACTCTCTGACACAGCTGTACTG 919
QY 955 tcttttaacttagagacatgagaaacctggcgataaaaaaggtatttacagttgagttctgc 1014
Db 918 ACGTTTACTCTTGAGATATGCAACCTGAGAGCAGACAACTTACTGTAGAGTTTGT 859
QY 1015 cctcaagaagaggttcaaatcactaaactgtgctactgttaacttactcgtcgtgagacacaa 1074
Db 858 CCGCTTAAACGTCGTCGTCGTACCAATATAGCAACGGTTCTTACTGTGGAGAGCATATA 799
QY 1075 tgtctgcaaatgaactacagttgttaactgagcctgtgtacaaagtaaatctctctggt 1134
Db 798 AATACAGCAAGCGTAACAACCTGTGATCAATGAGCCTTGCCTACAAAGTAACTATTGCA 739
QY 1135 gctgattggtcttcaagtatgtaaacctgtggagtagtactctatctcagtagtcaaatcctgga 1194
Db 738 GCAGATTGGTCTTATGCTTGTAAAGCCTGTAGAATAATGATGATCTCCGTTTCCAATCTCGA 679
QY 1195 gaactgttcttcatgattcgtgataccaagatacaactccctctggtgtacagtaetc 1254
Db 678 GATCTTGTGTTGGAGATGTCGCTGTTGAAGACACTCTTCTCCGAGGTCACAGTCTCT 619
QY 1255 gaagctcctggtgagagatctgctgtaataaaagtgtgtggtggtatataaagaaagtgc 1314
Db 618 GAAGCTCAGGAGCTCAAAATTTCTGTAATAAAGTAGTTTGGAGCTGTCAAAAGAACTCAAT 559
QY 1315 ccaggaaacccctccagtttaaaacttgaagagctcaagttcctcgtgaaagattcaca 1374
Db 558 CTGGAGAGTCTCTACAGTATAAAGTTCTAGTAAGAGCACAAACTCTCTGGACAACTTCA 499
QY 1375 aatcaagttcagtaactagtgagttcactcgcggaacatgatactatcttgcgcagaacac 1434
Db 498 AATTAATCTTCTGTGAAGAGCTGCTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439
QY 1435 acaacacattggaaggtctctcagctaccatgctgctgatttagacacaaatgactc 1494
Db 438 ACAACTTACTGGAAGAGAGTGTGCTACTCATATGTCGTAGTAGATATCTTGTGACCT 379
QY 1495 atctgttaggagaaaaactctatctglatctgtgttaactaacctggttctctctgaa 1554
```

Db 378 GTTGTGTAGGAGAAATACGTTTACCGGTATTTGTGTCCACCAACAGAGGTTCTGCAGAA 319  
Qy 1555 gatacaacgtatcttaatactgaagttctcaaaagaactcagccaaatagcttctca 1614  
Db 318 GATACAAATGTTCTTTAATGCTTTAAATTCCTTAAAGAACTGCAACCTGTATCCTCTCT 259  
Qy 1615 ggtccaaactaaaggacgattcaggttaataccggtgttttcgacgctttaccctaaactc 1674  
Db 258 GGACCAACTARAGGAACGATTACAGGCAATACAGTAGTAGTATTCGATTCTACCTAGATTA 199  
Qy 1675 ggttcaagaatctgaagttctttcttaccctgaaggtattgtcccgagatgct 1734  
Db 198 GGTCTTAAAGAAACGTAGAGTTTCTGTACATTGAAAGCAGTAGTACGCTGGAGATGCT 139  
Qy 1735 cgcggcgaagctattcttctctctgtacacactgactccacagctatccagacacacaaaat 1794  
Db 138 CGTGGGGAAGCGATTCTTTCTTCGATACATTGACTGTTCACGATTTCTGTATACAGAGAA 79  
Qy 1795 acccagctgtattaaattctaa 1816  
Db 78 ACACACATCTATTAATCTTTGA 57

## RESULT 8

US-10-007-693-21

; Sequence 21, Application US/10007693

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

; FILE REFERENCE: 210121.515C2

; CURRENT APPLICATION NUMBER: US/10/007,693

; CURRENT FILING DATE: 2001-12-05

; NUMBER OF SEQ ID NOS: 157

; SEQ ID NO 21

; LENGTH: 731

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis serovar E

US-10-007-693-21

Query Match 17.2%; Score 363.6; DB 37; Length 731;  
Best Local Similarity 75.8%; Pred. No. 3.9e-86;  
Matches 450; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 1223 aagatacactccctcttggtgttacagtaactcgaagctctctgtgagagatctctgta 1282  
Db 4 aagacactctctcccgagtcacagttctgaagctcgcagagctcaaatctctgtga 63  
Qy 1283 ataaagtgttggtcgattaaagaaatgtgcccgaggagaaacccctccagtttaacttg 1342  
Db 64 ataaagtgttggtcgattaaagaaatgtgcccgaggagaaacccctccagtttaacttg 123  
Qy 1343 tagtgaagctcaagtctctctggaagattcacaatacgaattgagtaactagtgagctta 1402  
Db 124 tagtaagagcacaactctctggaacattcacaataatgtgtgtggaagagctgctcg 183  
Qy 1403 actgcgaagactgtatctctgcgagaaacacaacacacatctggaaggtctctgagta 1462  
Db 184 actgtgactgtactctctgcgagaaacacaactactggaagaggtgctgta 243  
Qy 1463 cccatagtgcgtattagacacaaatgactctctgttaggagaaataactgtctctc 1522  
Db 244 ctcaatctgtcgttagatagactgtgacctgtgtgtgtaggagaaataactgtttacc 303  
Qy 1523 gtatctgttaactaacctgggtctctgctgaagataactaactatcttaactctgaagt 1582  
Db 304 gtatttgtccacacagaggtctctgcagagatacaaatgtttcttaagtcttaaat 363  
Qy 1583 tctcaagaactctcagcaatagcttctctcaggtccaaactaaaggaacgatttcagta 1642  
Db 364 tctcaagaactctcagcaatagcttctctcaggtccaaactaaaggaacgatttcagta 423

Qy 1643 ataccggtttgttgcgaogctttaccataacggttcctaaagaaatctgtagaagtttctg 1702  
Db 424 ataccggtttgttgcgaogctttaccataacggttcctaaagaaatctgtagaagtttctg 483  
Qy 1703 ttaccttgaaggtattgtctcccgagatgctcgcgcgaagctattcttctctgata 1762  
Db 484 taacattgaagcagttacagctgagatgctcgtgggaagcagattcttctctccgata 543  
Qy 1763 cactgactccacagttacagacacagaaataccacgctgtattaaattctaa 1816  
Db 544 cattgacttccagttctgtatagacagagaatacacacatctattaaattctga 597

## RESULT 9

US-09-410-568-63

; Sequence 63, Application US/09410568

; GENERAL INFORMATION:

; APPLICANT: Probst, Peter

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir

; APPLICANT: Fang, Hang

; APPLICANT: Jen, Shyian

; APPLICANT: Stromberg, Erica Jean

; APPLICANT: Enghart, Susan E.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; FILE REFERENCE: 210121.469C2

; CURRENT APPLICATION NUMBER: US/09/410,568

; CURRENT FILING DATE: 1999-10-01

; NUMBER OF SEQ ID NOS: 109

; SOFTWARE: fastseq for Windows Version 3.0

; SEQ ID NO 63

; LENGTH: 269

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-410-568-63

Query Match 8.5%; Score 179.4; DB 18; Length 269;  
Best Local Similarity 79.2%; Pred. No. 5.4e-37;  
Matches 213; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 470 acgtagaatttgcagtcctcgttcacgaatacgtactctagtagtctctaccctattg 529  
Db 1 atgttgaatacacaagctgttcttaaatatgctacggtagtagtctctaccctattg 60  
Qy 530 aaatctctgtatagcaaaaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 589  
Db 61 aaatctctgtatagcaaaaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120  
Qy 590 gcgaagctgaattcgttaagcagtcctccagaaacacacacacacacacacacacacacac 649  
Db 121 gtgaagcaggttctgtacagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 180  
Qy 650 tcggaaaaatcgatccctgggtgcaggagataaatgcaaaattactgtatgtgtgtgtgtgtgt 709  
Db 181 ttggaaaaatcgatccctgggtgcaggagataaatgcaaaattactgtatgtgtgtgtgtgtgt 240  
Qy 710 ctcttaaaagaagttgt 738  
Db 241 ctcttaaaagaagttgt 269

## RESULT 10

US-09-426-571-63

; Sequence 63, Application US/09426571

; GENERAL INFORMATION:

; APPLICANT: Probst, Peter

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir

; APPLICANT: Fang, Hang

```

1  APPLICANT: Ten, Shyian
2  APPLICANT: Stromberg, Erica Jean
3  APPLICANT: Enghart, Susan E.
4  TITLE OF INVENTION: COMPOUNDS AND
5  TITLE OF INVENTION: DIAGNOSIS OF
6  FILE REFERENCE: 469C3
7  CURRENT APPLICATION NUMBER: US/0
8  CURRENT FILING DATE: 1999-10-22
9  NUMBER OF SEQ ID NOS: 168
10 SOFTWARE: FastSeq for Windows Ver
11 SEQ ID NO 63
12 LENGTH: 269
13 TYPE: DNA
14 ORGANISM: Chlamydia trachomatis
15 US-09-426-571-63

```

Query Match	8.5%	Score 179.4	DB 18	Length 269
Best Local Similarity	79.2%	Prod. No. 5.4e-37		
Matches 213	Conservative 0	Mismatches 56	Indels 0	Gaps 0
QY	470	acgtagaattgccagtcgcttcacagaatcgcctactctagtagatctctaccctattg	529	
Db	1	atgtgaatcacacagctgtctcaaatatgtcgcgttagctctccctactctgtg	60	
QY	530	aaatctctgtatagcgaataaagattgtgtgatgttattacacaacagctacctt	599	
Db	61	aaatctactctacaggtgaagggattgtgtgtattctactcagcaattacct	120	
QY	590	gcgaagctgaattctgaagcagtgatcagcaaacactctcaagtgtatgggaattag	649	
Db	121	gtgaagcagagttctgacgcagtgatcagcgcaactctactgtctgtatgaagt	180	
QY	650	tctggaaaatcgatcgctcgggtgcaggagataaagtcaaaattactgtatgggtaaaac	709	
Db	181	tttggaaaattgcgcctctaggacaagcgcaaaagagtaaaattactgtatgggtaaaac	240	
QY	710	ctcttaagaagggtgtgtcttccacagct	738	
Db	241	ctcttaagaagggtgtgtcttccacagct	269	

```

RESULT 11
US-09-454-684-63
; Sequence 63, Application US/09454684
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skelky, Yasir
; APPLICANT: Filing, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND
; TITLE OF INVENTION: DIAGNOSTIC
; FILE REFERENCE: 210121.469C4
; CURRENT APPLICATION NUMBER: US/09/4-
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 303
; SOFTWARE: FASTSEQ for Windows Versi
; SEQ ID NO 63
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-454-684-63

```

	Query Match	8.5%	Score 179.4;	DB 18;	Length 269;
	Best Local Similarity	79.2%;	Pred. No. 5.4e-37;		
	Matches 213;	Conservative	Mismatches 56;	Indels	Gaps 0;
QY	470	acgtgaaattggccagtcgcgttcacgaatacgcctactgagatctctctacccattg	529		
Db	1	atgttgaaatcacacaaagtccttcctaataatgctacggtaggatctccctatccgttg	60		

[illegible]

```

RESULT 12
US-09-454-684A-63
? Sequence 63, Application US/09454684A
? GENERAL INFORMATION:
? APPLICANT: Probst, Peter
? APPLICANT: Bhatia, Ajay
? APPLICANT: Skeiky, Yasir
? APPLICANT: Fling, Steve
? APPLICANT: Malsonneuve, Jeff
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
? PREVENTION OF CHLAMYDIAL INFECTION
? FILE REFERENCE: 210121.469CA
? CURRENT APPLICATION NUMBER: US/09/454,684A
? CURRENT FILING DATE: 1999-12-03
? NUMBER OF SEQ ID NOS: 310
? SOFTWARE: FastSeq for Windows Version 3.0/4.0
? SEQ ID NO 63
? LENGTH: 269
? TYPE: DNA
? ORGANISM: Chlamydia trachomatis
US-09-454-684A-63

```

Query Match	8.5%	Score 179.4;	DB 18;	Length 269;
Best Local Similarity	79.2%;	pred. No. 5.4e-37;		
Matches 213;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps 0;
<hr/>				
Qy	470	acgtagaaattggccagtcgtctccagaataacgcgtactgtaggatccttaccctaatg	529	
Dd	1	atgtgaatacacacaagtgtctctaatagtcatcggtaggatcctctatcctgttg	60	
<hr/>				
Qy	530	aatatccttgtcatgagcaaaaagattgtgtgatgtttgattacacacagcgtacctt	589	
Dd	61	aaattctgtacaggtaaagggattgtgtgagtattcattatcgagcaattaccat	120	
<hr/>				
Qy	590	gcgaagctgaattcgttaagcagtgatccagaaacacactctcacagtgatggaaattag	649	
Dd	121	gtgaagcadagttcgttacgcagtgatccagcgacaaactctactgtcgtgtgaagtag	180	
<hr/>				
Qy	650	tctggaaatcgtatgcgctgggtgcaggagataaatgcaaaattactgtatgggtaaaaac	709	
Dd	181	tttggaaattgaccgcttaggacaagcgaaaagagtaaattactgtatgggtaaaaac	240	
<hr/>				
Qy	710	ctcttaaaagaaggtgtcgtctccacgct	738	
Dd	241	ctcttaaaagaaggtgtcgtctctccacgt	269	

```

RESULT 13
US-09-556-877-63
; Sequence 63, Application US/09556877
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve

```

APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556.877  
CURRENT FILING DATE: 2000-04-19  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FastSEQ for Windows Version 3.0/4.0  
SEQ ID NO 63  
LENGTH: 269  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
E-09-556-877-63

Query Match	8.5%	Score 179.4	DB 22	Length 269
Best Local Similarity	79.2%	Pred. No. 5.4e-37		
Matches 213	Conservative	0	Mismatches 56	Indels 0
Gaps				
QY	470	acgtagaaattgcagtcctccagaatacgtctactctaggatctcttaccctattg	529	
pb	1	atgttgaaatacacaaagcttctcctaataatgctacggtaggatctcctatcctgttg	60	
QY	530	aaatcctgtctataggcaaaaagatgtgttgatgttgattacacaaacagctacct	589	
pb	61	aaattactgtcacaggtaaaaggatgtgttgatgttactactcgcgaattacct	120	
QY	590	gcgaagctgaattcctaagcagtgatccagaaaacactctcacaagtgtcggaaattag	649	
pb	121	gtgaacagagttctgcacagtgatccacgcgacaactctactgctggttaagct	180	
QY	650	tctggaaaatcgcagctcctgggtgcaggagataaatgcaaaattactgtatgggtaaaac	709	
pb	181	tttggaaaattgaccgcttaggacaaggcgaaaagataaaattactgtatcgggtaaaac	240	
QY	710	ctcttaagaaggttctgcttcacagct	738	
pb	241	ctcttaagaaggttctgcttcacagct	269	

```

RESULT 14
US-09-598-419-63
; Sequence 63, Application US/09598419
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Skeiky, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: fastseq for Windows Version 3.0/4.0
; SEQ ID NO 63
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-63

```

	Query Match	8.5%	Score 179.4	DB 22	Length 269	
	Best local Similarity	79.2%	Pred. No. 5.4e-37			
	Matches 213	Conservative	0	Mismatches 56	Indels	Gaps 0
QY	470	acgtagaaattgcagctccggtccagaatacagctactatgagctctcttaccctattg	529			
Db	1	atgtgaaatcacacaagctgttccataatctacgtcagtgaggatccctatcctgttg	60			
QY	530	aaatcctgtataggcaaaaagattgttattgttggattacacaagcagctacctt	589			
Db	61	aaattactgtctacaggtaaaagggtgtgttgattcattactcagcaattacct	120			
QY	590	gcgaagctgaattcgaagcagtgatccagaaacaactcctacaagtgatgggaaattag	649			

	121	180
Db	g t g a a g c a g a g t t c g t a c g a g t g a t c c a g g a c a a c t c c t a c t g c g a t g g t a a g c t a g	
QY	t c t g a a a a t c a t c g c t t g g t g c a g a g a t a a a t g c a a a t t a c t g t a t c g g t a a a a c	
Db	t t g g a a a t t g a c c g c t t a g g a c a a g g c g a a a a g a a a a t t a c t g t a t g g g t a a a a c	
QY	c t c t t a a a a a g g t t g c t g c t t c a c a g t	738
Db	c t c t t a a a a a g g t t g c t g c t t c a c a g t	269

```

RESULT      15
US-09-620-412A-63
; Sequence 63, Application US/09620412A
;
; GENERAL INFORMATION:
;
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412A
; CURRENT FILING DATE: 2001-07-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 63
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-620-412A-63

```

	Query Match	8.5%	Score 179.4;	DB 24;	Length 269;	0;
	Best Local Similarity	79.2%;	Pred. No. 5.4e-37;			
	Matches 213;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps	
QY	470	acgtgagaatttgcacgtcccttccagaaatcacgtactgttaggtatctcttaccctattg	529			
Dbb	1	atgttgaatacacacaagctttcctaataatgctacggtaggatctccctatctctgtg	60			
QY	530	aaatccttgcctataggcgaataagattgtgttgattacacaaacagctacctt	589			
Dbb	61	aaattactgtctacaggtaaaaaggattgtgttgattactactcagcaattaccat	120			
QY	590	gogaagctgaattcgttaagcagtgatccagaaacaaactcctacaagtgatgggaattag	649			
Dbb	121	gtgaacagagtgcttcacgcagtgatccagcgacaactcctactgctgatgggtaagctag	180			
QY	650	tctggaaaatcgatgcctcgggtgcaggagataaaatgcacaaattactgtatgggtaaaaac	709			
Dbb	181	tttggaaaattgaccgcttaggacaaggcgcaaaagtgtaaaattactgtatgggtaaaaac	240			
QY	710	ctcttaaaagaaggttgccttccacagct	738			

search completed: May 25, 2002, 22:45:29  
job time: 15504 sec







```
QY 2079 gagaacgg 2086
Db 58 GTATCTG 51

RESULT 2
US-10-027-632-211458/c
; Sequence 211458, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 211458
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-211458

Query Match
Best Local Similarity 54.8%; Score 42; DB 6; Length 684;
Matches 103; Conservative 1; Mismatches 81; Indels 3; Gaps 1;

QY 1899 aaagcggttttttataagctcttcaatcgaatcgtagagttcttatacaaga 1958
Db 235 ACAGGCTTTGCTGATTTTGTACCTTTTCAACAAACTCTTTGTCATTGATCTTTA 176
QY 1959 tatatttaagttctgaaactcaagattattttaaagcccatcttttaggtatgt 2018
Db 175 TATKTTTAAAGTCTCTATTTTGTATGCTCTCTAACCATCATTTTCTTTGTGCT 116
QY 2019 aactaaatttttaagcttttccctagtgtaacctgtctttaggaactacactag 2078
Db 115 ACT--AAATTTGGGTTCAGTTTGTCTGTTTTTCTTGTCCTTGAGTGCAACATTGG 59
QY 2079 gagaacgg 2086
Db 58 GTATCTG 51

RESULT 3
US-09-540-209B-379/c
; Sequence 379, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 379
; LENGTH: 1227
; TYPE: DNA

; ORGANISM: B. fragilis
US-09-540-209B-379

Query Match
Best Local Similarity 2.0%; Score 41.4; DB 5; Length 1227;
Matches 75; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1851 gcttagtagtagctttcaagaagtagccgcttttagtacccttagtactactaaagcggtttt 1910
Db 251 GCTATAGAGCGCTTTATCAAAATAAGCCCGTCAGTGGTTGAAATACACAAGCTCACATC 192
QY 1911 ttgtttataagctcttcaatcccaatcgtagagttcttcttaacaagaatattattaaagt 1970
Db 191 TGGTAAACTAGGGTTATCAATCCCAATCTGAGCGAAACTTAAACGAAAGATCTTCTTTTAAAA 132
QY 1971 tctgaaatcc 1981
Db 131 AGAGGATATAC 121

RESULT 4
US-09-789-189-1188/c
; Sequence 1188, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789.189
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1188
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-789-189-1188

Query Match
Best Local Similarity 1.9%; Score 40.2; DB 5; Length 285;
Matches 111; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1848 tctgcttttagtagtagctttcaagaagtagccgcttttagtacccttagtactactaaagcggtt 1907
Db 256 TTTTCTTTTATAGAGAGCTAAAGAGAAATAAAAAATATATATTATTGTTGTTGAAATTTGTA 197
QY 1908 ttttcttttataagctcttcaatcccaatcgtagagttcttcttaacaagaatattatt 1967
Db 196 TTTTATTTTATAAAAGGGTAGAAGTTTAAAAAGTGGTAAAAATTTGTTATTATTAATTA 137
QY 1968 agttctgaaatcccaagaattatttttaaagcccatcttcttaggtatgtaataat 2027
Db 136 TTATTATTAATTAATGATATATTTAAAAAGAAATTTTAAAAATGAAATAATAATAA 77
QY 2028 ttttaattaaagcttttcttagtgtaacctgtcttcttaggaactacact 2076
Db 76 TGTAATAATTAGAAATTGTAATAATTTAAAAAGATGTTCTTTAGTAGAAAAAATT 28

RESULT 5
US-10-027-632-116847/c
; Sequence 116847, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116847
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1149)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-116847

Query Match
Best Local Similarity 1.9%; Score 39.2; DB 6; Length 1149;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 atcggtagttagagatgaataattctgactacattcaattcaataataaaccacaa 63
DB 807 ATATGGAATGTAGAATATATATCTTCAATTAATACTAAACAGCTATATCTTA 748
QY 64 tttgagggtgaagtttcaaacacattctaccgagtgagcagaagaaataaa 119
DB 747 TGTTTTGGTAATGTCATAGCCAAAATCTATTAAATGGATCAAGAAACAATAA 692

RESULT 6
US-10-027-632-134009/c
; Sequence 134009, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134009
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134009

Query Match
1.8%; Score 38.6; DB 6; Length 566;
US-10-027-632-157624
; Sequence 157624, Application US/10027632
```

```

Best Local Similarity 55.9%; Pred. No. 6.5;
Matches 71; Conservative 1; Mismatches 55; Indels 0; Gaps 0;

QY 1945 ttctaatcaagatattatttaagttcttgaaacttaagattattttaaagcccat 2004
DB 375 TTCATTATTAAATTTGTTCTGGATTTTGTAGATTGTTATCTTTATGTAACCCCTGAT 316
QY 2005 ctttttagtgatgaataataaatttttaataagcgttttcttaagtaacctgtctttt 2064
DB 315 GGTTTCAGGATATCATTTAAATTTGTTTCTTANATTTATTTTCCAGACTTCCTC 256
QY 2065 agdaact 2071
DB 255 AGGAATT 249

RESULT 7
US-10-027-632-257761/c
; Sequence 257761, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257761
; LENGTH: 629
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-257761

Query Match
1.8%; Score 38.4; DB 6; Length 629;
Best Local Similarity 52.5%; Pred. No. 7.5;
Matches 107; Conservative 0; Mismatches 96; Indels 1; Gaps 1;

QY 1835 agcgatattccgctctgttttagtagctttcaagaagaccgcttagacct-tac 1893
DB 534 AGACTTTCTCGTCTATATTTTTCATTGTTTAAAAAATGTTTTTTTCTCTGCTA 475
QY 1894 gtactaaagcgggtttttgttttataagctcttcaatccaatcgtagagttcttaac 1953
DB 474 ATATTAAGACTGTTATATTTTATTTAGTTTCATTAGGTCATGACATACATTGCTTTCAAATA 415
QY 1954 aagatattatttaagttcttctgaaatccatccatgaagattttttaaagcccatcttttagg 2013
DB 414 GCAAAACCTTGATCAGTTAACTGCAATTAATGACTTCTTTTAAAAAATATATAGTGGGTAGA 355
QY 2014 tatgtaataaataatttttaattaa 2037
DB 354 AATATAAGAAAAATATAAAAAATAA 331

RESULT 8
US-10-027-632-157624
; Sequence 157624, Application US/10027632
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157624
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-157624

Query Match          1.8%; Score 38.4; DB 6; Length 823;
Best Local Similarity 57.5%; Pred. No. 8;
Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1895 tactaaagcggtttttttgcttataagctcttcaatccatcgtagagttcttcaatca 1954
Db 413 tagcttggggatacatatttttaagtttttcaaggagaaaataatatttttcaaaa 472

QY 1955 aagatattattaaagttctcgaaatcctaagattatttttaaaagcccatcttttaggt 2014
Db 473 aaaaatttttaagttttcaaaaaccttataaaaaccttataaaaataatatttttaagt 532

RESULT 9
US-10-027-632-132469
; Sequence 132469, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132469
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Human
```

```
US-10-027-632-132469

Query Match          1.8%; Score 38.4; DB 6; Length 898;
Best Local Similarity 49.5%; Pred. No. 8.2;
Matches 96; Conservative 1; Mismatches 97; Indels 0; Gaps 0;

QY 1909 tttgtttttataagctcttcaatccatcgtagagttcttcaatcaaaagatatattttaa 1968
Db 86 ttttgggtcttattccctctcgcaataagaaaaatagtattattctcagctcatgtggtg 145

QY 1969 gtcttgaaatcctcaagatttttttaaaagcccatcttttttaggtatgttaataaaatt 2028
Db 146 attcaggtttatcggaattcttcaacaataagcccatcaatttttttttwaattccattt 205

QY 2029 ttttaataagcttttctctagtgtaacctgcttcttttaggaactacatcaggagaaacggt 2088
Db 206 ttcaattcaggttttcaatttttctctctctggttttgagaaactataaaaggtttgtaaca 265

QY 2089 tgtcatcaaatcta 2102
Db 266 tgccttaaatcta 279

RESULT 10
US-09-789-189-251
; Sequence 251, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lellias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789,189
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 251
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-789-189-251

Query Match          1.8%; Score 38; DB 5; Length 359;
Best Local Similarity 52.5%; Pred. No. 8.1;
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1880 ctttagtaccttaccgtactactaaagcggtttttttttataagctcttcaatccaatcgt 1939
Db 135 ccttaaccttattattttctctccacaaattatttcataatttacttttaaaattacaat 194

QY 1940 agagttcttcaatcaaaagatatatttttaagtttctgaaatcctcaagatttttttaaaag 1999
Db 195 taacttttttcttcacatatatttttttaatracttaacattcaacttttaaaatac 254

QY 2000 cccatcttttttaggtatgtaataaaatttttaattaa 2037
Db 255 attccttttttaaaatttcaattactatttttttta 292

RESULT 11
US-10-027-632-229235/c
; Sequence 229235, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 2000-04-30
; PRIOR FILING DATE: 2000-07-12
```

```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 229235
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-229235

```

```

Query Match      1.8%; Score 37.8; DB 6; Length 588;
Best Local Similarity 52.2%; Pred. No. 10;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1889 ctacgtactaagaagcggttttttttttataagctcttcaatccaatcgtagagttctt 1948
DB 436 CTGATGAACATATACGTTGCTTTCATTTGCAATTGCAATAGCAGAACACAGTACAATTCA 377

QY 1949 taatcaagatattattagttctgaatcctcaagattatttttaaaagcccatctt 2008
DB 376 TATTGAAGTTTAATATTGTATGATAGTAAATTTCTAGATTTTATTTAAATGTTATAT 317

QY 2009 ttaggatgtaattaaatttttaataagcttttccctagt 2049
DB 316 TTACTTTAATCATAGTAATTTAAACTTGCCTTCAAAATGT 276

```

```

RESULT 12
US-10-027-632-233567/c
; Sequence 233567, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233567
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-233567

```

```

Query Match      1.8%; Score 37; DB 6; Length 633;
Best Local Similarity 46.9%; Pred. No. 16;

```

```

Matches 112; Conservative 1; Mismatches 126; Indels 0; Gaps 0;

QY 1811 ttctaaggaattatcctaaagcagagcgatatctcgctctgtcttagagatagctttcaaa 1870
DB 325 TTGATGCAATTAAGAGTAATAATTAACCATGACACACACTACTCTCTCTAAGAAATCCCTTAA 266

QY 1871 gaagtaccgcttttagtacccttaagctactaaagcggtttttttttttataagctcttcaa 1930
DB 265 TAAATAAAATTTTATTATGTTGGTGTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 206

QY 1931 tccaatcgtagagttcttcaatcaagatatatttttaagtttctgaaatcctcaagattta 1990
DB 205 AACCATTTTGGAGTTAATAATATATATATAATATATATATATATATATATATATATATATAT 146

QY 1991 ttttaaaagcccatcttttttagtattgaatcaataaaatttttaataagctttcctagt 2049
DB 145 CTATCATCTCTCTCTGATGATAYATAATATTATACAAATCTCTTAAAACTTATTATTAT 87

```

```

RESULT 13
US-10-027-632-263505
; Sequence 263505, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263505
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-263505

```

```

Query Match      1.8%; Score 37; DB 6; Length 646;
Best Local Similarity 53.9%; Pred. No. 16;
Matches 76; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1906 tttttttttttataagctcttcaatccaatcgtagagttttcttaatacaagattatt 1965
DB 440 ttagtctttttttttccattccatcccttcattattttttttttttttttttttttttta 499

QY 1966 taagtcttcaaatcctaaagatttttttaaaagcccatcttttttaggtatgaattcaa 2025
DB 500 tttttatcattcttcaaaatttttttttttttttttttttttttttttttttttttttt 559

QY 2026 atttttaattaaagcttttct 2046
DB 560 ttttattgtattttcttctatct 580

```

```

RESULT 14
US-10-027-632-243685
; Sequence 243685, Application US/10027632
; GENERAL INFORMATION:

```

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
3  TITLE OF INVENTION: Polymorphisms in the Human Genome
4  FILE REFERENCE: 08827.123
5  CURRENT APPLICATION NUMBER: US/10/027,632
6  CURRENT FILING DATE: 2002-04-30
7  PRIOR APPLICATION NUMBER: US 60/218,006
8  PRIOR FILING DATE: 2000-07-12
9  PRIOR APPLICATION NUMBER: US 60/198,676
10 PRIOR FILING DATE: 2000-04-20
11 PRIOR APPLICATION NUMBER: US 60/193,483
12 PRIOR FILING DATE: 2000-03-29
13 PRIOR APPLICATION NUMBER: US 60/185,218
14 PRIOR FILING DATE: 2000-02-24
15 PRIOR APPLICATION NUMBER: US 60/167,363
16 PRIOR FILING DATE: 1999-11-23
17 PRIOR APPLICATION NUMBER: US 60/156,358
18 PRIOR FILING DATE: 1999-09-28
19 PRIOR APPLICATION NUMBER: US 60/146,002
20 PRIOR FILING DATE: 1999-08-09
21 NUMBER OF SEQ ID NOS: 325720
22 SOFTWARE: FastSeq for Windows Version 4.0
23 SEQ ID NO 243685
24 LENGTH: 598
25 TYPE: DNA
26 ORGANISM: Human
27 US-10-027-632-243685

```

Query Match	1.7%	Score 36.8;	DB 6;	Length 598;
Best Local Similarity	51.9%;	Pred. NO. 18;		
Matches 83;	Conservative 0;	Mismatches 77;	Indels 0;	Gaps 0;
QY 1906	ttttttgttttaagctcttcacatccatcgtagagttcttcaatcaagatattatt	1965		
Db 253	ttctattctctgccaatgcacatccatcgtagtaatttttgaagatgaaaa	312		
QY 1966	taagttctcgaactcctaagatttattttaaagccatctttttaggtagtgaattaaa	2025		
Db 313	taagtgctcaaatattattatttttgcacatatatttcaattttccagcagcattata	372		
QY 2026	atttttaataagcttccctcagtagtaacctgctcttta	2065		
Db 373	aattactctccactttaaattgaatttttggcccttttata	412		

```

RESULT      15
US-10-138-701-23
; Sequence 23, Application US/10138701
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Staphylococcus aureus genes and polypeptides
; FILE REFERENCE: PB484
; CURRENT APPLICATION NUMBER: US/10/138,701
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/512,255A
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/098,964
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-05
; PRIOR APPLICATION NUMBER: PCT/ US99/19726
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 2204
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-138-701-23

```

	Query Match	1.7%	Score 36.8;	DB 6;	Length 2204:
	Best Local Similarity	50.0%;	Pred. No. 25;		
Matches	92; Conservative	0;	Mismatches	92; Indels	0; Gaps
Qy	531	aatcctgtctatagcgcaaaaaagattggttgtagtggattaccacaacagactaccttg	590		
Db	1168	aatcatgcagagatatccgaattggtgatagtgtttagtgtaaaaacagagtgacat	1227		
Qy	591	cgaagctgaatctgaagcagtcgataccagaacaaactctcacaggatgctggaaatagt	650		
Db	1228	calactcgaagtgtacgttgtattcccagaacgtagacgtgaggatgctgcacatlcca	1287		
Qy	651	ctggaaaaatcgatcgctcgtgggtcgaggagataaattgcaaaattactgtatgggtgaaaaacc	710		
Db	1288	tatgccaacccccattgccaaagtttggaacatgaattagtcgtatttgaaaggcgaaagtayc	1347		
Qy	711	tcct	714		
Db	1348	acct	1351		

Search completed: May 25, 2002, 22:12:23  
Job time: 8192 sec



QY 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
Db 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300  
Db 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300  
QY 301 NVATVTCGHHKCSANVTTVVNEPCVOVNIISGADMSYCKPVEYSISVSNPGDLVLDVV 360  
Db 301 NVATVTCGHHKCSANVTTVVNEPCVOVNIISGADMSYCKPVEYSISVSNPGDLVLDVV 360  
QY 361 IQDTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLQFKLVKAQVPGRTNOVAVTSE 420  
Db 361 IQDTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLQFKLVKAQVPGRTNOVAVTSE 420  
QY 421 SNGCTCTCAETTTTHWGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480  
Db 421 SNGCTCTCAETTTTHWGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540  
Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHYV 556  
Db 541 DTLTSPVSDTENTHYV 556

## RESULT 2

US-09-523-647-2

; Sequence 2, Application US/09523647  
; GENERAL INFORMATION:  
; APPLICANT: MURDIN, ANDREW D.  
; APPLICANT: COMEN, RAYMOND P.  
; APPLICANT: WANG, JOE  
; APPLICANT: DUNN, PAMELA  
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING  
; FILE REFERENCE: 032931/0227  
; CURRENT APPLICATION NUMBER: 60/123,967  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: 60/123,966  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-523-647-2

Query Match 100.0%; Score 2951; DB 19; Length 556;

Best Local Similarity 100.0%; Pred. No. 3.6e-266;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKLRVTVTLTSMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVRVLR 60  
Db 1 MSLKLRVTVTLTSMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVRVLR 60  
QY 61 RNKQPEOKSRGAFCDKEFYPCEEGRCPQVEAQOQESCYRLYSKVNDNDCNVEICQSVPE 120  
Db 61 RNKQPEOKSRGAFCDKEFYPCEEGRCPQVEAQOQESCYRLYSKVNDNDCNVEICQSVPE 120  
QY 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
Db 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180

QY 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300  
Db 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300  
QY 301 NVATVTCGHHKCSANVTTVVNEPCVOVNIISGADMSYCKPVEYSISVSNPGDLVLDVV 360  
Db 301 NVATVTCGHHKCSANVTTVVNEPCVOVNIISGADMSYCKPVEYSISVSNPGDLVLDVV 360  
QY 361 IQDTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLQFKLVKAQVPGRTNOVAVTSE 420  
Db 361 IQDTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLQFKLVKAQVPGRTNOVAVTSE 420  
QY 421 SNGCTCTCAETTTTHWGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480  
Db 421 SNGCTCTCAETTTTHWGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540  
Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHYV 556  
Db 541 DTLTSPVSDTENTHYV 556

## RESULT 3

US-09-841-132-398

; Sequence 398, Application US/09841132  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 398  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-841-132-398

Query Match 100.0%; Score 2951; DB 22; Length 556;

Best Local Similarity 100.0%; Pred. No. 3.6e-266;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKLRVTVTLTSMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVRVLR 60  
Db 1 MSLKLRVTVTLTSMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVRVLR 60  
QY 61 RNKQPEOKSRGAFCDKEFYPCEEGRCPQVEAQOQESCYRLYSKVNDNDCNVEICQSVPE 120  
Db 61 RNKQPEOKSRGAFCDKEFYPCEEGRCPQVEAQOQESCYRLYSKVNDNDCNVEICQSVPE 120  
QY 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
Db 121 YATVGSPPYPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOBGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300  
Db 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300



Db 241 EVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Qy 301 NVATVTCGGHKCSANVTTVNNEPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
Db 301 NVATVTCGGHKCSANVTTVNNEPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
Qy 361 IODTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETLOFLVVKAAQVPGRTNOVAVTSE 420  
Db 361 IODTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETLOFLVVKAAQVPGRTNOVAVTSE 420  
Qy 421 SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENVTYRICVTVNRGSAEDTNVSLIL 480  
Db 421 SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENVTYRICVTVNRGSAEDTNVSLIL 480  
Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTCLKGIAPGDARGEAILSS 540  
Db 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTCLKGIAPGDARGEAILSS 540  
Qy 541 DTLTSPVSDTENTHY 556  
Db 541 DTLTSPVSDTENTHY 556  
RESULT 4  
US-09-438-185-559  
; Sequence 559, Application US/09438185  
; GENERAL INFORMATION:  
; APPLICANT: Stephens, Richard  
; APPLICANT: Mitchell, Wayne  
; APPLICANT: Kalman, Sue  
; APPLICANT: Davis, Ronald  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
; FILE REFERENCE: 018941-000411US  
; CURRENT APPLICATION NUMBER: US/09/438.185  
; CURRENT FILING DATE: 1999-11-11  
; PRIOR APPLICATION NUMBER: US 60/108,279  
; PRIOR FILING DATE: 1998-11-12  
; PRIOR APPLICATION NUMBER: US 60/128,606  
; PRIOR FILING DATE: 1999-04-08  
; NUMBER OF SEQ ID NOS: 1074  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 559  
; TYPE: PR1  
; ORGANISM: Chlamydia pneumoniae  
US-09-438-185-559

Query Match 100.0%; Score 2951; DB 18; Length 559;  
Best Local Similarity 100.0%; Pred. No. 3.6e-266;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITIKIVASAETKPAPVPMATAKRVLR 60  
Db 4 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITIKIVASAETKPAPVPMATAKRVLR 63  
Qy 61 RNKQPVQKSRGAFCDKEYPCEEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120  
Db 64 RNKQPVQKSRGAFCDKEYPCEEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 123  
Qy 121 YATVGSYPYIEILAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180  
Db 124 YATVGSYPYIEILAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 183  
Qy 181 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKQEGPDCACLRCPVCYKI 240  
Db 184 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKQEGPDCACLRCPVCYKI 243  
Qy 241 EVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 244 EVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

Qy 301 NVATVTCGGHKCSANVTTVNNEPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
Db 304 NVATVTCGGHKCSANVTTVNNEPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 363  
Qy 361 IODTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETLOFLVVKAAQVPGRTNOVAVTSE 420  
Db 364 IODTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETLOFLVVKAAQVPGRTNOVAVTSE 423  
Qy 421 SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENVTYRICVTVNRGSAEDTNVSLIL 480  
Db 424 SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENVTYRICVTVNRGSAEDTNVSLIL 483  
Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTCLKGIAPGDARGEAILSS 540  
Db 484 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTCLKGIAPGDARGEAILSS 543  
Qy 541 DTLTSPVSDTENTHY 556  
Db 544 DTLTSPVSDTENTHY 559  
RESULT 5  
US-09-438-185A-559  
; Sequence 559, Application US/09438185A  
; GENERAL INFORMATION:  
; APPLICANT: Stephens, Richard  
; APPLICANT: Mitchell, Wayne  
; APPLICANT: Kalman, Sue  
; APPLICANT: Davis, Ronald  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
; FILE REFERENCE: 018941-000411US  
; CURRENT APPLICATION NUMBER: US/09/438.185A  
; CURRENT FILING DATE: 2002-03-13  
; PRIOR APPLICATION NUMBER: US 60/108,279  
; PRIOR FILING DATE: 1998-11-12  
; PRIOR APPLICATION NUMBER: US 60/128,606  
; PRIOR FILING DATE: 1999-04-08  
; NUMBER OF SEQ ID NOS: 1074  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 559  
; TYPE: PR1  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; OTHER INFORMATION: CPn0557  
US-09-438-185A-559

Query Match 100.0%; Score 2951; DB 18; Length 559;  
Best Local Similarity 100.0%; Pred. No. 3.6e-266;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITIKIVASAETKPAPVPMATAKRVLR 60  
Db 4 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITIKIVASAETKPAPVPMATAKRVLR 63  
Qy 61 RNKQPVQKSRGAFCDKEYPCEEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120  
Db 64 RNKQPVQKSRGAFCDKEYPCEEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 123  
Qy 121 YATVGSYPYIEILAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180  
Db 124 YATVGSYPYIEILAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 183  
Qy 181 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKQEGPDCACLRCPVCYKI 240  
Db 184 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKQEGPDCACLRCPVCYKI 243  
Qy 241 EVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 244 EVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

```
QY 301 NVATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 304 NVATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLVLHDVV 363
QY 361 IQDTLPSSGVTVLEAPGGEICCNKVMRIKEMCPGETTQFKLVKAQVPGPFTNOVAVTSE 420
Db 364 IQDTLPSSGVTVLEAPGGEICCNKVMRIKEMCPGETTQFKLVKAQVPGPFTNOVAVTSE 423
QY 421 SNGCTCTCAEATTHMKGLAATHMVCVLTNDPVCVNTVNTGNSAEDTNNVSLIL 480
Db 424 SNGCTCTCAEATTHMKGLAATHMVCVLTNDPVCVNTVNTGNSAEDTNNVSLIL 483
QY 481 KFSKELQPTASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEATLSS 540
Db 484 KFSKELQPTASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEATLSS 543
QY 541 DTLTSPVSDTENTHVV 556
Db 544 DTLTSPVSDTENTHVV 559

RESULT 6
US-09-841-132-441
; Sequence 441, Application US/09841132
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 441
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-441
```

```
Query Match 73.2%; Score 2160.5; DB 22; Length 553;
Best Local Similarity 71.7%; Pred. No. 2.7e-192;
Matches 401; Conservative 75; Mismatches 68; Indels 15; Gaps 4;

QY 1 MSKLIRRVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMATAKVKRLVR 60
Db 7 MNKLIRRAVTIFAIVTSVASFASGLVETSMASLSLNTNVLADTKAK--DNTSHKSKAR 64
QY 61 RN---KOPVEOKSRGAFCDKEFYPCBEGRCQPVQAEQSCYGRLYSVKVNDDCNVEICQS 117
Db 65 KNSKETPVNR-----KQVAPVHESKA--TGPQKDSFCGRMYTVKVNDDRNVEITQA 114
QY 118 VPEYATVGSPPYPIELAIKGGKDCVDTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRL 177
Db 115 VPEYATVGSPPYPIELAIKGGKDCVDTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRL 174
QY 178 GAGDKCKITVWVKPLKEGCGCTTAATVCACPELRSYTKCGQPAICIKQEGPCDCLRCPCV 237
Db 175 GQGEKSKITVWVKPLKEGCGCTTAATVCACPELRSYTKCGQPAICIKQEGPCDCLRCPCV 234
QY 238 YKIEVW--NTGSAIARNTVDNPPDGYSHASQORVLSNGLDMRPGDKKVFTEFCPCOR 297
Db 235 YKINIYNOGTATARNVNVNPPDGYSHASQORVLSNGLDMRPGDKKVFTEFCPCOR 294
QY 298 QITNATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLVLH 357
Db 295 RATNIATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLVLH 354
QY 358 DVVIQDTLPSSGVTVLEAPGGEICCNKVMRIKEMCPGETTQFKLVKAQVPGPFTNOVAV 417
Db 355 DVVVEDTLPSSGVTVLEAPGGEICCNKVMRIKEMCPGETTQFKLVKAQVPGPFTNOVAV 414
```

```
QY 418 TSNCGCTCTCAEATTHMKGLAATHMVCVLTNDPVCVNTVNTGNSAEDTNNV 477
Db 415 KSCSDCGCTCTCAEATTHMKGLAATHMVCVLTNDPVCVNTVNTGNSAEDTNNV 474
QY 478 LILKFSKELQPTASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAT 537
Db 475 LMLKFSKELQPTASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAT 534
QY 538 LSSDTLTPVSDTENTHVV 556
Db 535 LSSDTLTPVSDTENTHVV 553

RESULT 7
US-09-201-228A-1036
; Sequence 1036, Application US/09201228A
; GENERAL INFORMATION:
; APPLICANT: Griffiths, Remy
; APPLICANT: Holseth, Susan K.
; APPLICANT: Zagursky, Robert John
; APPLICANT: Metcalfe, Benjamin J.
; APPLICANT: Peek, Joel A.
; APPLICANT: Sankaran, Banumathi
; APPLICANT: Fletcher, Leah Diane
; TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE
; TITLE OF INVENTION: AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN
; FILE REFERENCE: 9710-0004-999
; CURRENT APPLICATION NUMBER: US/09/201.228A
; CURRENT FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/107,077
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: FR 97-16034
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: FR 97-15041
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 5981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1036
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-201-228A-1036
```

```
Query Match 71.3%; Score 2104.5; DB 16; Length 557;
Best Local Similarity 70.2%; Pred. No. 4.7e-187;
Matches 393; Conservative 76; Mismatches 73; Indels 19; Gaps 6;

QY 1 MSKLIRRVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMATAKVKRLVR 60
Db 7 MNKLIRRAVTIFAIVTSVASFASGLVETSMASLSLNTNVLADTKAK--DNTSHKSKAR 64
QY 61 RN---KOPVEOKSRGAFCDKEFYPCBEGRCQPVQAEQSCYGRLYSVKVNDDCNVEICQS 117
Db 65 KNSKETPVNR-----KQVAPVHESKA--TGPQKDSFCGRMYTVKVNDDRNVEITQA 114
QY 118 VPEYATVGSPPYPIELAIKGGKDCVDTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRL 177
Db 115 VPEYATVGSPPYPIELAIKGGKDCVDTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRL 174
QY 178 GAGDKCKITVWVKPLKEGCGCTTAATVCACPELRSYTKCGQPAICIKQEGPCDCLRCPCV 237
Db 175 GQGEKSKITVWVKPLKEGCGCTTAATVCACPELRSYTKCGQPAICIKQEGPCDCLRCPCV 234
QY 238 YKIEVW--NTGSAIARNTVDNPPDGYSHASQORVLSNGLDMRPGDKKVFTEFCPCOR 295
Db 235 YKINIYNOGTATARNVNVNPPDGYSHASQORVLSNGLDMRPGDKKVFTEFCPCOR 294
QY 296 RGQITNATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLV 355
Db 295 RGQITNATVTCGGHKCSANVTTVNPPCQVNNISGADMSYVCKPVEYSISVSNPGDLV 355
```

Db 295 RGRATNTAMVYCGGHKNTASVTTVINPCQVSIAGADWSYVCKPVEYVISVNSPGDLV 354  
Qy 356 LHDVVIQDTPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFLVKV--AQVPGRFN 413  
Db 355 LRDVVVKTLPSPGVTVLEAAGASICNKKVVWTVKELNPGESLOYKVLVARGAQTPOGFN 414  
Qy 414 QVAVTSNCGTCTSCAETTHMKGLAATHMCVLDITNDPICVGENTVYRICVNRGSAED 473  
Db 415 NVVVKSCDGTCTSCAETTVKGVAAATHMCVVDTCDPVCVGENTVYRICVNRGSAED 474  
Qy 474 TNVSLILKFSKELOPTASSGPKTGTISNTVVFDALPKLGSKEVSEFVTLKGIAPGDAR 533  
Db 475 TNVSLMLKFSKELOPVSFSGPKGTITNTVVFDLPLRLSKETVEFVTLKAVSAGDAR 534  
Qy 534 GEALLSSDITLSPVSDTENTHY 556  
Db 535 GEALLSSDITLSPVSDTENTHY 557

RESULT 8  
US-10-020-269-38  
; Sequence 38, Application US/10020269  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, William M.  
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/10/020,269  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 38  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-10-020-269-38

Query Match 4.8%; Score 141; DB 24; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 60 RNNKQPVQKSGAFCDKEFPCEE 84  
Db 1 RNNKQPVQKSGAFCDKEFPCEE 25

RESULT 9  
US-09-902-540-10065  
; Sequence 10065, Application US/09902540  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15845)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 10065  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-10065

Query Match 4.5%; Score 132; DB 23; Length 902;  
Best Local Similarity 22.8%; Pred. No. 0.022;  
Matches 104; Conservative 40; Mismatches 161; Indels 152; Gaps 23;  
Qy 103 SYKVNDQC-NVETICO-----SVPEYATVGSPIETILAIKGGKDCVDVVIQOLPCEA 153  
Db 517 TVTVDGAGNASTCQFTVTVRDITAP---TIVCPAEVVEAMGPAGAVAGFV---LPQAT 570  
Qy 154 EFVSSDPETT--PTSDGKL-----VWKIDRLGAGDKCKITVWVKPLKEGCCCTAATV 203  
Db 571 DAVTASPOVTASPSGSEFMGLSTPVTLTATDDAGNVGSCQFLVTVRD-----TAAPV 623  
Qy 204 CACPRLSRVTRCGQPAICIQBDPCACLRCPVYKIEVNTGSAIARNVTVDNBPQGY 263  
Db 624 LTCPE-----EVIVETAGGATVSPAGQAVDVV-----SOQPVY 660  
Qy 264 SHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQTNVATVYCGGHKCSANVTVVNE 323  
Db 661 SHASG---ADPFLG;-----GTLVSV-TATDAAGNAAACQFTVTVRD 697  
Qy 324 P-----CVQ-VNISGADMSYVCKPVEYSISVSNPGDLVLDVVIQDTPSGVTVLEAPG 376  
Db 698 TTPPELGCPODVVLEAD-----TQCAVVVTAGVQPRDLVTHDPSIAFSOTSG----- 745  
Qy 377 GEICCNKVVWRIKEMCPGETLQFLVKVKAQVPGRF---TNOAVTSESNCGTCTCAET 433  
Db 746 -----SRFPLGTTAVAVTASDAAGNEASCRAV 773  
Qy 434 THWKGUATHMCVLD-----TNDPICVGENTVYRICVNRGSAEDTNVSLILKFSKELOPI 489  
Db 774 TVQDTPAPQCMPSDVSVETQDP---EGTVLSYA---PASAVDGVSSVTVVAYS-----L 821  
Qy 490 ASSGPKGTISGNT-VVFDALPKLGSKEVSEFVSVTLK 525  
Db 822 ASGG---HFASGTTVPVVVTTATDTAGNAAQCFLVSVR 855  
RESULT 10  
US-07-808-458-2  
; Sequence 2, Application US/07808458  
; GENERAL INFORMATION:  
; APPLICANT: Fleming, Robert J.  
; TITLE OF INVENTION: Artavanis-Tsakonas, Spyridon  
; TITLE OF INVENTION: Nucleotide And Protein Sequences Of The  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07/808,458  
; FILING DATE: 19911211  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-005  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 8698864/9741  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1404 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-07-808-458-2

Query Match 4.4%; Score 130.5; DB 3; Length 1404;  
Best Local Similarity 20.5%; Pred. No. 0.058;  
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;  
Qy 14 LTMASCFASGGIEAAVAESLTKIVASAETKPAVPMTAKKVLRRNKPVEQKSRGA 73  
Db 543 LTTTATAIGSNLSALLAALTSAVASTSLAIGPC-INAKECR-----NQP-----GS 590  
Qy 74 FDCKEFYPCBEGRCQVPEAQOESCGRLYSVKVNDCNVEICQSVPEYATVGSPIEL 133  
Db 591 FA-----CICKEG-----WGGVTCENLDDC-VGQCRN----- 617  
Qy 134 AIGKKDCVDVITQOLPCEAEFVSDPE-----TTPTS DG-----KLVWK-----IDRLG- 178  
Db 618 ---GATCIDLVNDYRCACASGFTGRDCTDIDECATSPCRNGGECVDMVGKFCNICPLGY 674  
Qy 179 AGDKC---KITVWVKPLKGGCFTA--ATVCACPELRSYTKCQO-PAIC----- 221  
Db 675 SGLCEAEKENCTPSPCLEGHCLNTPEGYCHCPDRAGKHCEQLRPLCSQPPCNEGCPA 734  
Qy 222 -----IKQEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250  
Db 735 NVSLATSATTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHGTGTCF 793  
Qy 251 RNVTVDNVPV-----DG-----YSHASQORVLS-----FNLGDMRPG-- 282  
Db 794 HNLNECSPNCRNGGICLDGDDFTCECMGWTGKRCERATGCGYAGOCNGGTCMFGAP 853  
Qy 283 DKV-----FTVEFCPQRRGQITNVATVYCGGHKCSANVTTVNNEPCVQVNSGA 333  
Db 854 DKALQPHRCAPGWTGLCAE-----AIDQCRGCPCHNGGT-----CE-----SCA 894  
Qy 334 DW-SYVCKPVEYSISVNSPGDLVLDVVDLPSGVTVLEAPGECICCNKVVWRKEMC 392  
Db 895 GWFRVC-----AQGSGPDCRINNECSPPQCGGATCIDGIGYS-----IC 939  
Qy 393 PGETLQKLVVKAQVPGRTNOVATVSESNGCTCSAET-----TTHWKGAAATH 443  
Db 940 P-----PGRHGLRCEILLSDPKSACQNASNTISPYTALNRSQNWLDIALTG 985  
Qy 504 VVFDAL 509  
Db 1036 VCVPAL 1041

RESULT 11  
US-09-195-524-2  
; Sequence 2, Application US/09195524  
; GENERAL INFORMATION:  
; APPLICANT: Ish-Horowitz, David  
; APPLICANT: Henrique, Domingos M.P.  
; APPLICANT: Lewis, Julian H.  
; APPLICANT: Myat, Anna M.  
; APPLICANT: Fleming, Robert J.  
; APPLICANT: Artavanis-Tsakonas, Spyridon  
; APPLICANT: Mann, Robert S.  
; APPLICANT: Gray, Grace E.  
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/195,524  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/611,729  
; FILING DATE: 06-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-037  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1404 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-195-524-2

Query Match 4.4%; Score 130.5; DB 15; Length 1404;  
Best Local Similarity 20.5%; Pred. No. 0.058;  
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;  
Qy 14 LTMASCFASGGIEAAVAESLTKIVASAETKPAVPMTAKKVLRRNKPVEQKSRGA 73  
Db 543 LTTTATAIGSNLSALLAALTSAVASTSLAIGPC-INAKECR-----NQP-----GS 590  
Qy 74 FDCKEFYPCBEGRCQVPEAQOESCGRLYSVKVNDCNVEICQSVPEYATVGSPIEL 133  
Db 591 FA-----CICKEG-----WGGVTCENLDDC-VGQCRN----- 617  
Qy 134 AIGKKDCVDVITQOLPCEAEFVSDPE-----TTPTS DG-----KLVWK-----IDRLG- 178  
Db 618 ---GATCIDLVNDYRCACASGFTGRDCTDIDECATSPCRNGGECVDMVGKFCNICPLGY 674  
Qy 179 AGDKC---KITVWVKPLKGGCFTA--ATVCACPELRSYTKCQO-PAIC----- 221  
Db 675 SGLCEAEKENCTPSPCLEGHCLNTPEGYCHCPDRAGKHCEQLRPLCSQPPCNEGCPA 734  
Qy 222 -----IKQEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250  
Db 735 NVSLATSATTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHGTGTCF 793  
Qy 251 RNVTVDNVPV-----DG-----YSHASQORVLS-----FNLGDMRPG-- 282  
Db 794 HNLNECSPNCRNGGICLDGDDFTCECMGWTGKRCERATGCGYAGOCNGGTCMFGAP 853  
Qy 283 DKV-----FTVEFCPQRRGQITNVATVYCGGHKCSANVTTVNNEPCVQVNSGA 333  
Db 854 DKALQPHRCAPGWTGLCAE-----AIDQCRGCPCHNGGT-----CE-----SCA 894  
Qy 334 DW-SYVCKPVEYSISVNSPGDLVLDVVDLPSGVTVLEAPGECICCNKVVWRKEMC 392  
Db 895 GWFRVC-----AQGSGPDCRINNECSPPQCGGATCIDGIGYS-----IC 939  
Qy 393 PGETLQKLVVKAQVPGRTNOVATVSESNGCTCSAET-----TTHWKGAAATH 443  
Db 940 P-----PGRHGLRCEILLSDPKSACQNASNTISPYTALNRSQNWLDIALTG 985

QY 444 MCVLDTNDPICGENTVYVICVTRNGSAEDTNVSLILKFSKELQPIASSGPTKGTISGNT 503  
Db 986 RTEDDENACVCEN-----GTSRCTNLWCGLPNCYKVDPLSKSSNLGVCQKQHE 1035  
QY 504 VVFDAL 509  
Db 1036 VCVPAL 1041

## RESULT 12

US-09-944-849-8  
; Sequence 8, Application US/09944849  
; GENERAL INFORMATION:  
; APPLICANT: Nickoloff, Brian  
; APPLICANT: Miele, Lucio  
; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMEN  
; TITLE OF INVENTION: MALIGNANT AND BENIGN SKIN DISORDERS BY MODULATING THE NOTCH PATH  
; FILE REFERENCE: 212583  
; CURRENT APPLICATION NUMBER: US/09/944,849  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/229,614  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 1404  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-944-849-8

Query Match 4.4%; Score 130.5; DB 23; Length 1404;  
Best Local Similarity 20.5%; Pred. No. 0.058;  
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;

QY 14 LTSMAFCASGAGIEAAVAESLITKIVASAEKPAVPMTAKKRVLRNRKQVQEKSRGA 73  
Db 543 LTTATATIGSLNSTALLAALTSVASTSLAIGPC-INAKECR-----NOP-----GS 590  
QY 74 FCDKEFYPCCEGRCPQVEAQECYGRLYSVKVNDDCNVEICQSVPEATVGSPPYIEIL 133  
Db 591 FA---CICKEG-----WGGVTCANLDDC-VQCRN----- 617  
QY 134 AIGKDCVDVITQOLPCEAEFVSSDPE-----TPTSDG-----KLWVK-----IDRLG- 178  
Db 618 ---GATCIDLVNDYRCACASGFTGRDCEIDECATSPCRNGECVDMVGKFNCTCPUGY 674  
QY 179 AGDKC---KITVWKPLKEGCCFTA--ATVCACPELRSYTKCGQ-PAIC----- 221  
Db 675 SGLCEBEAKENCTSPCLGHCCLNTPEGYXCHCPDPDRAGKHCEQLRPLCSQPPCNEGCPA 734  
QY 222 -----IKQGPDC-----ACLRCPV---CYKIEVWNTGSAIA 250  
Db 735 NVSLATSATTTTTTTTATTTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHTGTCPE 793  
QY 251 RNYTVDNRPV-----DG-----YSHASGORVLS-----FNLGDMRPG-- 282  
Db 794 HNLNECSPNCRNGICLDGSDGDTFCQMSGWTGKRCSERATGCGYAGQCGQGTGCMCPAP 853  
QY 283 DKVK-----FTVEFCPPRRGQITNVAITYCGGHKCSANVTTVNNEPCVQVNISSGA 333  
Db 854 DKALQPHRCAPGWTGLFEAE-----AIDCRGQCPCHNGT-----CE-----SGA 894  
QY 334 DW-SYCKPVEYSISVNPGLVLHDVVVITQTLPSGVTVLEAPGGICCNKVVWRIKEMC 392  
Db 895 GWFRCVC-----AQGFSPDCRINNVNCSQPQCGGATCIDGGIGVSC-----IC 939  
QY 393 PGETLOFLVWKAQVPGRFTNQAVTSESNGCTCSAET-----TTHWKGLAATH 443  
Db 940 P-----PGRHGLURCEILLSDPKSACQNASNTISPYTALNRSQNWLDIATG 985

QY 444 MCVLDTNDPICGENTVYRICVTNRNGSAEDTNVSLILKFSKELQPIASSGPTKGTISGNT 503

Db 986 RTEDDENACVCEN-----GTSRCTNLWCGLPNCYKVDPLSKSSNLGVCQKQHE 1035  
QY 504 VVFDAL 509  
Db 1036 VCVPAL 1041

## RESULT 13

PCT-US01-08631-46695  
; Sequence 46695, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 46695  
; LENGTH: 1483  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (76)..(121)  
; OTHER INFORMATION: kv TRANSCRIPTASE REVERSE II ORF2 domain identified by  
; OTHER INFORMATION: eMATRIX, accession number DM01354N, p-value=1.000e-40, raw sco  
; OTHER INFORMATION: 13.17  
; NAME/KEY: DOMAIN  
; LOCATION: (783)..(1314)  
; OTHER INFORMATION: Immunoglobulin domain identified by Pfam, accession name ig,  
; OTHER INFORMATION: E-value=1.9e-38, Pfam score of 128.9  
PCT-US01-08631-46695

Query Match 4.4%; Score 130.5; DB 1; Length 1483;  
Best Local Similarity 19.3%; Pred. No. 0.063;  
Matches 118; Conservative 90; Mismatches 217; Indels 185; Gaps 29;

QY 35 ITKIVASAEKTPAPV-----PWTAKKYR-----LVRRNKOPVEQKSRGACDKE--- 78  
Db 626 VSQVVGCTTHPRDLRFNRIRIQCAFRRRLNLTLLNNQ-IKRIPSGAFEDLENLK 584  
QY 79 FYPCGEGRCPQVEAQECYGRLYSVKVNDD---CNVEI---COSVPEYATVGS----- 126  
Db 685 YLYLYKNEIQSIDRQAFKGLASLEQLRLDSNTLHCDCEILWLADLLATYAESGNAQAAAI 744  
QY 127 -PYPIELAIKGRKDCVDVITQOLPCEAEFVSSDPEPTPTSDGKLWVKIDRLGAGDKCKI 185  
Db 745 CEYPRRI---QGRSVATITPEELNCERPRITSEPDADVTSGNTVYFTCRABGNPKPEI 800  
QY 186 TVWV--RPLKEGCCFTAATVCACPELRSYTKGQPAICIKQEGPDCACLRCP-VCYKIEV 242  
Db 801 -IWRNKSLSERHAADTGLSSRVDL-----TCLPNKG---SLLCPQTCGCPVK 846  
QY 243 VNTGSAIARNVVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVTFECPQRRGQITNV 302  
Db 847 AKTALAAAVRVTEGSPADHY-----MRPGEPLHSLTLC---RLALPNI 888  
QY 303 -----ATVT-----YCGGHKCSANVTTVNNEPCVQVNISSGADMSYVC---KPVEYSISV 349  
Db 889 QSHLQATLTCLWLAVELTGTISVASNVTSLL-----FTCFCKIKGTKKVLQVG 933  
QY 350 NPGDL-----VLHD--VVIQDTLPSPVTVLEA-----PGGEICCNKVVWRIKEMCPGET 396  
Db 934 NELSMTKDSRLNLLDGGTLMIQNTQETDQGIYQCMKNAVAGEVKTQEV-----T 982  
QY 387 LQFKLVVKAQVPGRFT-NQAVTSESNGCTCSAETTTTHWKGLAATHMVCVLTQNDP---- 452

```
Db 983 LRY-----FCSPGSLGNRVCTSSVSVCATSLQPNTEVLVGSVTLCSATGHPPRI 1037
Qy 453 -----ICVGENTVYRICVTNRGSAEDTNVSLILKFSK 484
Db 1038 SWTRGDRTPLPVDPVRNVTSPGSLYIQNVVQDGEVACSATNNIDSVHATAFIIV--Q 1094
Qy 485 ELQPIASSGPTKGTISGNVTFD-----ALPKLGSKESVEFSVTLKGIAPGDAR 533
Db 1095 ALPQFTVTPQDRVVIIEGOTVDFQCEAKGNPPVPIAWTKGGSQSV-----R 1141
Qy 534 GEAILSSDTL 543
Db 1142 RHLVLSGTL 1151

RESULT 14
US-60-173-464-28067
; Sequence 28067, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: LI, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173.464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28067
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-28067

Query Match 4.4%; Score 129.5; DB 26; Length 756;
Best Local Similarity 21.0%; Pred. No. 0.029;
Matches 117; Conservative 60; Mismatches 216; Indels 163; Gaps 30;

Qy 5 IRRVVTVLALTSMAFCFASGGIEAAVAESLTKIVASAETKPAP-----VPMATAKK 55
Db 81 IRRVEAIWSESCYCFHFKICIQRWANDSMQMKVKAEOQNGGHHNLGFEVPPKQK 140
Qy 56 -----VRLVRNRKQPVQKSR-GAFCDKEFYPCBEGRCQPVQEAQOQESYGRLYSVKVVND 109
Db 141 SLHWCCPKRRDYQADPKPTQYNCFCGKEVNPENQPLVP-HSCGEHC-GKLLQPKCGHD 198
Qy 110 CNVEICQSV-----EYATV-----GSPYPI-----EILAIGKKDCVDV- 143
Db 199 CKL-LCHPGCPPCAQAQVCLCGKSSPRVRCIDKQWTCQOTCKELLACGKHKCNQVC 257
Qy 144 -----VITQOL-PCBAEFVSSDPETPTSDGKLWVKIDRL-GAGDKCKITVWVK--- 190
Db 258 HQPKCPPTCKSLQPCQ-----RESKVNCSDRK--WKQNVCGAPFACGLHICEKVCH 312
Qy 191 --PLKEGCCFTAAATVACAPCLRSYTKCGQPAICIKQBPDCACLRCPVCYKI--EYVNTG 246
Db 313 AGPCGDGECPLQVRSCPC-----GKNCCGDCQPCCKICGKOLSCN 353
Qy 247 SAIRANVTVDNP-----VDPGYSHASGQVLSFNLGDMRP----- 281
Db 354 KHKQSVCHNGPCYCKLESQINRCRGTKKSPVCGRERSA--RIVCLELCTRITPKCHHA 411
Qy 282 -----GDKVTFTEFCQRGQITNVATVTCGGHKCSA--NVTTVVNEPCQVNVISG 332
Db 412 IKHCHKGD-----CSP-CGQVCGLPNDTSGKHICAKRCHAEAVRNKP--KEARPOA 461
Qy 333 ADMSYVCKPVEYSISVNPGLDLVLDVVIQDTLPSTGVTLEAPGGEI-----CCNKYVWRI 388
Db 462 KKEYKALP-----HPRCEGVIVTCIGGHEVATWPCWNSKPTSC 501
Qy 389 KEMCPGETLOFKL-VYKAQVPGREFTNQVAVTSSESNGT-CTSCAETTTTHWKLAAATHMVCV 446
```

```
Db 502 QRS---ARQLKCGNHKSLVCHP---VPLPDMSAQTGCANCEEGCTVPRPTGCIHACP 555
Qy 447 LQNDPICVGENTVYR 462
Db 556 KCHPPPCAPCNFVIK 571

RESULT 15
US-09-614-150-36948
; Sequence 36948, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36948
; LENGTH: 757
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-36948

Query Match 4.4%; Score 129.5; DB 20; Length 757;
Best Local Similarity 21.0%; Pred. No. 0.029;
Matches 117; Conservative 60; Mismatches 216; Indels 163; Gaps 30;

Qy 5 IRRVVTVLALTSMAFCFASGGIEAAVAESLTKIVASAETKPAP-----VPMATAKK 55
Db 82 IRRVEAIWSESCYCFHFKICIQRWANDSMQMKVKAEOQNGGHHNLGFEVPPKQK 141
Qy 56 -----VRLVRNRKQPVQKSR-GAFCDKEFYPCBEGRCQPVQEAQOQESYGRLYSVKVVND 109
Db 142 SLHWCCPKRRDYQADPKPTQYNCFCGKEVNPENQPLVP-HSCGEHC-GKLLQPKCGHD 199
Qy 110 CNVEICQSV-----EYATV-----GSPYPI-----EILAIGKKDCVDV- 143
Db 200 CKL-LCHPGCPPCAQAQVCLCGKSSPRVRCIDKQWTCQOTCKELLACGKHKCNQVC 258
Qy 144 -----VITQOL-PCBAEFVSSDPETPTSDGKLWVKIDRL-GAGDKCKITVWVK--- 190
Db 259 HQPKCPPTCKSLQPCQ-----RESKVNCSDRK--WKQNVCGAPFACGLHICEKVCH 313
Qy 191 --PLKEGCCFTAAATVACAPCLRSYTKCGQPAICIKQBPDCACLRCPVCYKI--EYVNTG 246
Db 314 AGPCGDGECPLQVRSCPC-----GKNCCGDCQPCCKICGKOLSCN 354
Qy 247 SAIRANVTVDNP-----VDPGYSHASGQVLSFNLGDMRP----- 281
Db 355 KHKQSVCHNGPCYCKLESQINRCRGTKKSPVCGRERSA--RIVCLELCTRITPKCHHA 412
```

QY 282 -----GDKKVTVEFCORRGQITNATVTYCGHKCSA--NVTTVVNEPCVQVNISG 332  
DB 413 IKHCHKGD-----CPP-CGQVGLPNDTSKCGHICKARCHEAVRVNRP-KEARPOA 462  
QY 333 ADMSYVCKPVEYSISVSNFGDLVLDHVDVJQDTLPSTVLEAPGGEI-----CONKVVWRI 388  
DB 463 KKEYKALP-----HPRCEGVIVTCIGGHEVATWPCWNSKPTSC 502  
QY 389 KEMCPGETLOFKL-VVKAQVPGFTNOAVTSESNGT-CTSCAETTHWKGLAATHMCV 446  
DB 503 QRSK---ARQLKCGNHKCSLVCHF---VPLPQDMSAQTCANCEGCTVPRPTGCIHACP 556  
QY 447 LDTNDPICVGENTVYR 462  
DB 557 KGCHPPPCAPCNFVIK 572

Search completed: May 25, 2002, 22:20:57  
Job time: 509 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:14:33 ; Search time 64.71 Seconds  
(without alignments)  
243,232 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 2951  
Sequence: 1 MSKLIRRVTVLALTSASC.....ILSSDTLTPVSDPENTHYV 556

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 116914 seqs, 28308587 residues

Total number of hits satisfying chosen parameters: 116914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA-New.\*

1: /cgn2.6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2.6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	120.5	4.1	716	5	US-09-312-283B-183	Sequence 183, App
2	118.5	4.0	1529	5	US-09-312-283B-396	Sequence 396, App
3	114.5	3.9	1036	6	US-10-121-049-142	Sequence 142, App
4	114.5	3.9	1036	6	US-10-121-050-142	Sequence 142, App
5	114.5	3.9	1036	6	US-10-121-053-142	Sequence 142, App
6	114.5	3.9	1036	6	US-10-121-043-142	Sequence 142, App
7	114.5	3.9	1036	6	US-10-121-044-142	Sequence 142, App
8	114.5	3.9	1036	6	US-10-121-047-142	Sequence 142, App
9	114.5	3.9	1036	6	US-10-121-054-142	Sequence 142, App
10	114.5	3.9	1036	6	US-10-121-055-142	Sequence 142, App
11	114.5	3.9	1036	6	US-10-121-057-142	Sequence 142, App
12	114.5	3.9	1036	6	US-10-121-058-142	Sequence 142, App
13	114.5	3.9	1036	6	US-10-121-060-142	Sequence 142, App
14	114.5	3.9	1036	6	US-10-121-063-142	Sequence 142, App
15	114.5	3.9	1036	6	US-10-123-108-142	Sequence 142, App
16	114.5	3.9	1036	6	US-10-123-154-142	Sequence 142, App
17	114.5	3.9	1036	6	US-10-123-156-142	Sequence 142, App
18	114.5	3.9	1036	6	US-10-123-157-142	Sequence 142, App
19	114.5	3.9	1036	6	US-10-123-212-142	Sequence 142, App
20	114.5	3.9	1036	6	US-10-123-213-142	Sequence 142, App
21	114.5	3.9	1036	6	US-10-123-109-142	Sequence 142, App
22	114.5	3.9	1036	6	US-10-121-041-142	Sequence 142, App
23	114.5	3.9	1036	6	US-10-121-045-142	Sequence 142, App
24	114.5	3.9	1036	6	US-10-121-046-142	Sequence 142, App
25	114.5	3.9	1036	6	US-10-121-051-142	Sequence 142, App
26	114.5	3.9	1036	6	US-10-121-040-142	Sequence 142, App

27	114.5	3.9	1036	6	US-10-121-048-142	Sequence 142, App
28	114.5	3.9	1036	6	US-10-121-052-142	Sequence 142, App
29	114.5	3.9	1036	6	US-10-121-061-142	Sequence 142, App
30	114.5	3.9	1036	6	US-10-121-042-142	Sequence 142, App
31	114.5	3.9	1036	6	US-10-121-055-142	Sequence 142, App
32	114.5	3.9	1036	6	US-10-121-059-142	Sequence 142, App
33	114.5	3.9	1036	6	US-10-124-822-142	Sequence 142, App
34	114.5	3.9	1036	6	US-10-123-903-142	Sequence 142, App
35	114.5	3.9	1036	6	US-10-124-817-142	Sequence 142, App
36	114.5	3.9	1036	6	US-10-124-819-142	Sequence 142, App
37	114.5	3.9	1036	6	US-10-124-823-142	Sequence 142, App
38	114.5	3.9	1036	6	US-10-125-704-142	Sequence 142, App
39	114.5	3.9	1036	6	US-10-123-215-142	Sequence 142, App
40	114.5	3.9	1036	6	US-10-123-235-142	Sequence 142, App
41	114.5	3.9	1036	6	US-10-123-236-142	Sequence 142, App
42	114.5	3.9	1036	6	US-10-123-261-142	Sequence 142, App
43	114.5	3.9	1036	6	US-10-123-322-142	Sequence 142, App
44	114.5	3.9	1036	6	US-10-124-820-142	Sequence 142, App
45	114.5	3.9	1036	6	US-10-123-771-142	Sequence 142, App

## ALIGNMENTS

## RESULT 1

US-09-312-283B-183

; Sequence 183, Application US/09312283B

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, James D.

; APPLICANT: Sleeman, Matthew

; APPLICANT: Oerust, Rene

; APPLICANT: Murlison, James G.

; APPLICANT: Kumble, Krishanend D. Isolated from Skin Cells

; TITLE OF INVENTION: Compositions and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312.283B

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 183

; LENGTH: 716

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283B-183

Query Match 4.1%; Score 120.5; DB 5; Length 716;  
Best Local Similarity 20.4%; Pred. No. 0.17; 126; Indels 191; Gaps 24;  
Matches 92; Conservative 42; Mismatches 126;

QY	72	GAFCDKKEFYPCPEGRCQ-----PVEAQOESC---YGRLY-----SVKVNDD	109
DB	313	GEHCIDFDDCDNKCKNGAHCCTDAVNGYTCVCEPGYSGLFCFSPPMVFLRTSPCDNFD	372
QY	110	C-----NVEICQSVPEYATGSPVPEILAIGKKDC-----VDVWITQOLPCEAE	154
DB	373	CQGAOCILIRVNEPICQCLPGY-----LGEK-CEKLVSVSILVNKESYLQIP	418
QY	155	FVSSDPEF-----TPTSDGKLVWKIDELGAGDKCKITVWVKPLKGGCCFATACACP	207
DB	419	SAKVRPOTNTLQIATDESGILLYK-----GDKHIAV---ESIEGI-----	458
QY	208	ELRSYTKCGOPATCIKQEGPDCACLRCPVCYKIEVNTGS-----AIARNVTVDNP	258
DB	459	-RASDYTGSHFASAI-----YSVETINDGNFHVIELITLSSLSLSDVGG	502
QY	259	VPDGYSHASQRYLSEF-----LGDMPGDKKVFTEFCPQRGQITNVATVTCGGHKCS	314
DB	503	SPKIITNLKQSTLNFDSPLYVGGM-PGKNNVASLRQAPGQNG-----TSFHCCI	551
QY	315	ANVTTVNEPCVQVNSIGADWSVCKPYEYSVSNPGDLVLHDVWIODTLPSGVTVLEA	374

Db 552 RN-----LVINSELQDRKV--PMOTGI-----L 573  
Qy 375 PGGEICCNKVVWRIKEMCPGETLOFLVVKVKAQVPGRTNQAVTSBNGCTCTSCAETTT 434  
Db 574 PGGEPCHKV-----CAHGTCQ-----PSSQSGFTCECE--- 603  
Qy 435 HWKGLAATHMCVLDTNDPTCVGENTVYRICV 465  
Db 604 GWMG----PLCDQRTNDP-CLGNKCVHGTCL 629  
RESULT 2  
US-09-312-283B-396  
: Sequence 396, Application US/09312283B  
: GENERAL INFORMATION:  
: APPLICANT: Watson, James D.  
: APPLICANT: Strachan, Lorna  
: APPLICANT: Sleeman, Matthew  
: APPLICANT: Onrust, Rene  
: APPLICANT: Murison, James G.  
: APPLICANT: Kumble, Krishanand D.  
: TITLE OF INVENTION: Compositions Isolated from Skin Cells  
: FILE REFERENCE: 11000.1011c2  
: CURRENT APPLICATION NUMBER: US/09/312.283B  
: CURRENT FILING DATE: 1999-05-14  
: NUMBER OF SEQ ID NOS: 425  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 396  
: LENGTH: 1529  
: TYPE: PRT  
: ORGANISM: Mouse  
US-09-312-283B-396

Query Match 4.0%; Score 118.5; DB 5; Length 1529;  
Best Local Similarity 20.3%; Pred. No. 0.6;  
Matches 94; Conservative 39; Mismatches 115; Indels 215; Gaps 25;  
Qy 72 GAFCDKEFYPCBGRQ-----PVEAQESC---YGRLY-----SVKVNDD 109  
Db 1070 GEHCDIDFDCQNKCKNGAHCTDAVNGYTCVCEGYSGLFCEFPMPVLLRTSPCDNFD 1129  
Qy 110 C-----NVEICQSVPEYATVGSPPYPIEILAIKDKCDVDVITQQLPCEAEFVSSD 159  
Db 1130 CONGAOCIIIRVNEPICOLPGY-----LGEK-CEKLV-----SVNEVKE 1168  
Qy 160 -----PFT-----TPTSDGKLVKIDRLGAGDKCKITVWVKPLKESCCFTAA 201  
Db 1169 SYLQIPSAKYRPTNITLOIATDEDSGILLYK-----GKDHIAV----- 1208  
Qy 202 TVCACPEL-----RSYTKGGOPATCIKQEGPDCACLRCPVCYKIEVVNTGS----- 247  
Db 1209 -----ELYGRVRASYDTGSHPSAI-----YSVETINDGNFHIVELLT 1247  
Qy 248 -AIARNYVDNPVPGYSHASQORVLSFN-----LGDMPRGDKKVTVEFCPQRRGQITNV 302  
Db 1248 LDSSLSLSDVGGSPKTIITNLRSKSTLNFDSPLYVGGM-PGKNNVASLRQAPGONG----- 1301  
Qy 303 ATVTYCGCHKCSANVTTVVNEPCVOVNISGADWSYVCKPVEYSISVSNPGDLVLDHWIQ 362  
Db 1302 -----TSFHCICRN-----LYINSELQDRKV--PMOTGI----- 1329  
Qy 363 DTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLOFLVVKVKAQVPGRTNQAVTSBNG 422  
Db 1330 -----LPGCEPCHKV-----CAHGTCQ-----PSSQ 1351  
Qy 423 CGTCTCAETTHWKGLAATHMCVLDTNDPTCVGENTVYRICV 465  
Db 1352 SGFTCECE---GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1386

RESULT 3  
US-10-121-049-142  
: Sequence 142, Application US/10121049  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: Deforge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: APPLICANT: Zhang, Zemin  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P3330R1C17  
: CURRENT APPLICATION NUMBER: US/10/121.049  
: CURRENT FILING DATE: 2002-04-12  
: Prior Application removed - See File Wrapper or Palm  
: NUMBER OF SEQ ID NOS: 550  
: SEQ ID NO 142  
: LENGTH: 1036  
: TYPE: PRT  
: ORGANISM: Homo Sapien  
US-10-121-049-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;  
Best Local Similarity 19.2%; Pred. No. 0.75;  
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;  
Qy 75 CDKEFYPCBGRQVPEAQESQYGRLYSVK-----VNDDCNWTQ 116  
Db 376 CERYVVP--EGECCPV-----CEDPVYPNPNFAGCYANGLILAHGDRWREDDCT--FCQ 425  
Qy 117 SVPEYATVGSPPYPIEILAIKDKCDVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167  
Db 426 CVN-----GERHCVATVCGTCTNPKVPGECCPVCEERTIITVDPAC 469  
Qy 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEG---C-----FTAAT 202  
Db 470 GELSNCITLTGKDCINGFKRDHNG---CRTCOCINTTEELCSEKQGCITLNCPPGFLTDAQ 525  
Qy 203 VCACPELRSYTKGGOPATCIK-----QEGPD-CACLRCP--VCYKIEVVNTGSAIA 250  
Db 526 NCIEICRPRPKRPIICDKYCPGLLKNKHGDCICRCKCPELSCSI----- 575  
Qy 251 RNVVDNPVPGYSHASQORVLSFNLDMPRGDKKVTVEFCPQRRGQITNVATVYTCGG 310  
Db 576 -----CPLGFQDOSHGCCLICKRE 594  
Qy 311 HKCSANVTTVVNEPCVQV---NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356  
Db 595 ASASAG-PPILSGTCLTVDGHGHHKNEESWHDGRCYCLNGREMCALITCPVPACGNPTI 653  
Qy 357 H-----DVVLOTLPSPGVTVLEAPGGEICCNKVVWRI-----KE 390  
Db 654 HPGCCPSCADDFVQKRPSTPSICHAPGGEYFVEGETWNIDSTQCTCHSGRVLCETE 713  
Qy 391 MCPGETLOFLVVKVKAQVPGRTNQAVTSBNGCTCTSCAETTHWKGLAATHMCVLDTN 450  
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRPSLRNNSVNYCNDEG 758  
Qy 451 DPICVGENTVYRICV 466

```
Db 759 DIFLAESWKPVDCTS 774

RESULT 4
US-10-121-050-142
: Sequence 142, Application US/10121050
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3330R1C20
: CURRENT APPLICATION NUMBER: US/10/121.050
: CURRENT FILING DATE: 2002-04-12
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 142
: LENGTH: 1036
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-121-050-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

Qy 75 CDKEFYPCERGCQCPVEAQOESCGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVP--EGECCPV-----CEDPVYFPNPNAGCYANGLILAHGDRWRDDCT--FCQ 425
Qy 117 SVPEYATVGSYPPIELATGKKDCVDVVITO-----OLPCEAEFYSSDP---ETTPSD 167
Db 426 CVN-----GERHCYATVCGQCTNPNVKVFGECPCVEEPTIITVDPPAC 469
Qy 168 GKL-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C-----FTAA 202
Db 470 GELSNCTLTGKDCINGFKRDHNG-----CRTCCINTELCSERKQGTCLNCPFGFLTDAQ 525
Qy 203 VCACPELRSYTKGQPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLGKLNKHGDCICRCKKCPELSCKSI----- 575
Qy 251 RNVTVDNPPVDGYSHASGQVLSFNLDGMRPGDKKVFTEFCPPRRGQITNVTATYCGG 310
Db 576 -----CPLGFOODSHGCLICKRE 594
Qy 311 HKCSANVTTVVNEPCVOV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG--PPILSGTCLTVGHHKNEESHDGCRECYCLNGREMCALITCPVACGNPTI 653
Qy 357 H-----DVIQDTLPSTVLEAPGGEICCNKVVRI-----KE 390
Db 654 HPGQCCPSCADDFWQKPELSTPSICHAPGGEYFVEGETWNIDSCOTCHSGRVLCE 713
Qy 391 MCPGETLQFLVVAQVPGFRFTNOVAVTSNCGTCTSCAETHHKGALAAHMCVLDTN 450
Db 714 VCP-----PLLCQNPSSR-----TQDSCCPQCTDQPPRPSLRNNSVFNKNDG 758
```

```
QY 391 MCPGETLQFKLVKAQVPGRFTNOAVTSESNGTCTSCAETTHHKGLAATHMCVLDTN 450
Db 714 VCP-----PILCONPSR-----TODSCCPQCTDQPPRPSLRNNSVPNYCKNDEG 758
QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPVDVCTS 774

RESULT 6
US-10-121-043-142
; Sequence 142, Application US/10121043
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C15
; CURRENT APPLICATION NUMBER: US/10/121,043
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-043-143
```

```
Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
```

```
QY 75 CDKEFYPCCEGRCPVEAQOQESCYGRLYSVK-----VNDGCVNVEICQ 116
Db 376 CERYVYP--EGECCPV-----CEDPVYFPNPNAGCYANGLILAHGDRWREDDCT--FCQ 425
QY 117 SVPEYATVGSYPPIETILAIKGDVWVITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCYATVCGQTCTNPVKVPGECPCVCEPTIITVDPPAC 469
QY 168 GKLIV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAAT 202
Db 470 GELSNCLTGKDCINGFKRDHNG-----CRTOCQINTEELCSERKQCTLNCPPGFLFLDAQ 525
QY 203 VCACPELRSYTKCGQPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLLLKNKHGDCICRCKCPELSCKSI-----575
QY 251 RNVTVDPNPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQDSDSHGCLICKRE 594
QY 311 HKCSANVTTVNNEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356
Db 595 ASASAG--PPLISGTCCLTVGDHHHKNESWHDGRCYCLNGREMCALITCPVACGNPTI 653
QY 357 H-----DVIQDTLPSTGVTYVLEAPGGEICCNKVVWRI-----KE 390
```

```
Db 654 HPQCCPCSCADDEVQKPELSTPSCICHAPGGEYFVEGETWNIDSCQTCTCHSCRVLCETE 713
QY 391 MCPGETLQFKLVKAQVPGRFTNOAVTSESNGTCTSCAETTHHKGLAATHMCVLDTN 450
Db 714 VCP-----PILCONPSR-----TODSCCPQCTDQPPRPSLRNNSVPNYCKNDEG 758
QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPVDVCTS 774

RESULT 7
US-10-121-044-142
; Sequence 142, Application US/10121044
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C15
; CURRENT APPLICATION NUMBER: US/10/121,044
; CURRENT FILING DATE: 2002-04-11
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-044-142
```

```
Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
```

```
QY 75 CDKEFYPCCEGRCPVEAQOQESCYGRLYSVK-----VNDGCVNVEICQ 116
Db 376 CERYVYP--EGECCPV-----CEDPVYFPNPNAGCYANGLILAHGDRWREDDCT--FCQ 425
QY 117 SVPEYATVGSYPPIETILAIKGDVWVITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCYATVCGQTCTNPVKVPGECPCVCEPTIITVDPPAC 469
QY 168 GKLIV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAAT 202
Db 470 GELSNCLTGKDCINGFKRDHNG-----CRTOCQINTEELCSERKQCTLNCPPGFLFLDAQ 525
QY 203 VCACPELRSYTKCGQPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLLLKNKHGDCICRCKCPELSCKSI-----575
QY 251 RNVTVDPNPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQDSDSHGCLICKRE 594
QY 311 HKCSANVTTVNNEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356
```

Db 595 ASASAG-PPILSGTCLTVGHHHKNESWHDGCRCYCLNGREMCALITCPVPACGNPTI 653  
QY 357 H-----DVIQDTLPSSGVTVLEAPGGEICCNKVVWRI-----KE 390  
Db 654 HPGCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIDSCQTCTCHSGRVLCE 713  
QY 391 MCPGETLQFLVKAQVPGRETNOVAVTSBNGCTCTSCAETTHWKGLAATHMCLVDTN 450  
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRLSRNNSVNYCKNDEG 758  
QY 451 DPICVGVNTVYRICVT 466  
Db 759 DIFLAESWKPDDVCTS 774

RESULT 8  
US-10-121-047-142

; Sequence 142, Application US/10121047  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C4  
; CURRENT APPLICATION NUMBER: US/10/121.047  
; CURRENT FILING DATE: 2002-04-11  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 142  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-047-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;  
Best Local Similarity 19.2%; Pred. No. 0.75;  
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

QY 75 CDKEFYPCBGRCPVQAQESCYGRLYSVK-----VNDCCNVEICQ 116  
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLIILAHGDRWREDCT--FCQ 425  
QY 117 SVPEYATVGSPPYIEILAIKGDVVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167  
Db 426 CVN-----GERHCYATVCGTCTNPVKVFGCCPCVEEPTIITVDPPAC 469  
QY 168 GKLKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202  
Db 470 GELSNCTLTGKDCINGFKRDHNG---CRTQCINTBELCSERKQGLTNCPPFGLTDAQ 525  
QY 203 VCACPELRSYTKCGQPAICIK-----QEGPD-CACLRCP--VCYKIEVVNTGSAIA 250  
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKHGDCICRCKKCPELSCSKI----- 575  
QY 251 RNVTVDNPVDSHGSGQVRLSFLNGLDMRPGDKKVTVEPCQRRGQITNATVTCGG 310  
Db 576 -----CPLGFOQDSHGCLICKCRE 594

QY 311 HKCSANVTTVWNEPCVQV-----NISGADWSVCKPV-----EYSISVSNP-----GDLVL 356  
Db 595 ASASAG-PPILSGTCLTVGHHHKNESWHDGCRCYCLNGREMCALITCPVPACGNPTI 653  
QY 357 H-----DVIQDTLPSSGVTVLEAPGGEICCNKVVWRI-----KE 390  
Db 654 HPGCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIDSCQTCTCHSGRVLCE 713  
QY 391 MCPGETLQFLVKAQVPGRETNOVAVTSBNGCTCTSCAETTHWKGLAATHMCLVDTN 450  
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRLSRNNSVNYCKNDEG 758  
QY 451 DPICVGVNTVYRICVT 466  
Db 759 DIFLAESWKPDDVCTS 774

RESULT 9

US-10-121-054-142  
; Sequence 142, Application US/10121054  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C4  
; CURRENT APPLICATION NUMBER: US/10/121.054  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 142  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-054-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;  
Best Local Similarity 19.2%; Pred. No. 0.75;  
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

QY 75 CDKEFYPCBGRCPVQAQESCYGRLYSVK-----VNDCCNVEICQ 116  
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLIILAHGDRWREDCT--FCQ 425  
QY 117 SVPEYATVGSPPYIEILAIKGDVVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167  
Db 426 CVN-----GERHCYATVCGTCTNPVKVFGCCPCVEEPTIITVDPPAC 469  
QY 168 GKLKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202  
Db 470 GELSNCTLTGKDCINGFKRDHNG---CRTQCINTBELCSERKQGLTNCPPFGLTDAQ 525  
QY 203 VCACPELRSYTKCGQPAICIK-----QEGPD-CACLRCP--VCYKIEVVNTGSAIA 250  
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKHGDCICRCKKCPELSCSKI----- 575

```

QY 251 RNVTVNDPVPDGYSHASQORVLSFNLGDMRPGDKVFTVEFCPPORRGOITNVAITYCGG 310
Db 576 -----CPLGFGQDQSHGCLICKRE 594

QY 311 HKCSANVTTVVNEPCVOV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHHKNESWHDGCRCYCLNGREMCALITCPVPACGNPTI 653
QY 357 H-----DVTIQTLPSCVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPQCCPCSCADDFVOKPELSTPSICHAPGGEYFVEGETWNIDSTQCTCHSGRVLCE 713
QY 391 MCPGETLQFLKLVVKAQVGRFTNQAVTSESNGTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCQNSR-----TQDSCCPQCTDQFPFRPSLSRNNVSPNYCKNDEG 758

QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAAESWKPDVCTS 774

```

```

RESULT 10
US-10-121-056-142
; Sequence 142, Application US/10121056
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC23
; CURRENT APPLICATION NUMBER: US/10/121,056
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-056-142

```

```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

QY 75 CDKEFYPCBEGRCQPVFAQOQESCYGRLYSVK-----VNDDCNVEICQ 116
Db 376 CERYYVP--EGECCPV-----CEDVPYFPNNPAGCYANGLILAHGDRWREDDCT--FCQ 425
QY 117 SVPEYATVGSPIPIELAIKGGKDCVDVVITQ-----QLPCBAEFVSSDP---ETTPSTD 167
Db 426 CVN-----GERHCVATVCGOTCTNPVKVPGGCCPVCEPTIITVDPPAC 469
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202
Db 470 GELSNCITLTKDCINGEVRDRNG-----CRTQCQINTEELCSRKQKQCTLNCFPGLTDAQ 525
QY 203 VCACPELSYTKCGOPATCIK-----QEGPD-CACLRCP--VCYKIEWVNTGSAIA 250

```

```

Db 526 NCEICECRPRPKCRPIICDKYCPGLLLKNKHGCDICRCKKCPELSCSKI----- 575
QY 251 RNVTVNDPVPDGYSHASQORVLSFNLGDMRPGDKVFTVEFCPPORRGOITNVAITYCGG 310
Db 576 -----CPLGFGQDQSHGCLICKRE 594
QY 311 HKCSANVTTVVNEPCVOV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHHKNESWHDGCRCYCLNGREMCALITCPVPACGNPTI 653
QY 357 H-----DVTIQTLPSCVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPQCCPCSCADDFVOKPELSTPSICHAPGGEYFVEGETWNIDSTQCTCHSGRVLCE 713
QY 391 MCPGETLQFLKLVVKAQVGRFTNQAVTSESNGTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCQNSR-----TQDSCCPQCTDQFPFRPSLSRNNVSPNYCKNDEG 758

QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAAESWKPDVCTS 774

```

```

RESULT 11
US-10-121-057-142
; Sequence 142, Application US/10121057
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC13
; CURRENT APPLICATION NUMBER: US/10/121,057
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-057-142

```

```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

QY 75 CDKEFYPCBEGRCQPVFAQOQESCYGRLYSVK-----VNDDCNVEICQ 116
Db 376 CERYYVP--EGECCPV-----CEDVPYFPNNPAGCYANGLILAHGDRWREDDCT--FCQ 425
QY 117 SVPEYATVGSPIPIELAIKGGKDCVDVVITQ-----QLPCBAEFVSSDP---ETTPSTD 167
Db 426 CVN-----GERHCVATVCGOTCTNPVKVPGGCCPVCEPTIITVDPPAC 469
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202

```

```
Db 470 GELSNCTLTCKDCINGFKRDHNG-----CRTQCINTFELCSERKQGTLCNCPFGFLTDAQ 525
Qy 203 VCACPELRSYTKGCPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKRCPIICDKYCPGLLKNKHGCDICRCKKPELSCSKI-----575
Qy 251 RNVTVDPVDPGYSHASGORVLSFNLDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPPCVQV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHDGCRECYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIODTLPSGVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGCCPSCADDFVVKPELSTPSICHAPGGEYFVEGETWNIDSTCTCTCHSGRVLCE 713
Qy 391 MCPGETLOFLVYKAQVGRFTNQAVTSENCGTCTSCAETTHWKGLAATHMVCVLDN 450
Db 714 VCP-----PLLCQNPFR-----TQDSCCPQCTDQPPRPSLRNNSVNPYCKNDEG 758
Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDCVCTS 774

RESULT 12
US-10-121-058-142
Sequence 142, Application US/10121058
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C12
CURRENT APPLICATION NUMBER: US/10/121,058
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 142
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-058-142
```

```
Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
Qy 75 CKEFEYPCBEGRCQPVAEQAQESCYRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYIYP--EGECCPV-----CEDPVYFPNPNAGCYANGLILAHGRWRDEDDCT--FCQ 425
Qy 117 SVPEYATVGSPIYETILATGKDCVDVITQ-----QLPCAEEFVSDDP---ETPTSD 167
Db 426 CVN-----GERHCYATVCGQTCTNPNKVPKGECCPVCEETITVDPPAC 469
```

```
Qy 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC-----C-----FTAAT 202
Db 470 GELSNCTLTCKDCINGFKRDHNG-----CRTQCINTFELCSERKQGTLCNCPFGFLTDAQ 525
Qy 203 VCACPELRSYTKGCPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKRCPIICDKYCPGLLKNKHGCDICRCKKPELSCSKI-----575
Qy 251 RNVTVDPVDPGYSHASGORVLSFNLDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPPCVQV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHDGCRECYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIODTLPSGVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGCCPSCADDFVVKPELSTPSICHAPGGEYFVEGETWNIDSTCTCTCHSGRVLCE 713
Qy 391 MCPGETLOFLVYKAQVGRFTNQAVTSENCGTCTSCAETTHWKGLAATHMVCVLDN 450
Db 714 VCP-----PLLCQNPFR-----TQDSCCPQCTDQPPRPSLRNNSVNPYCKNDEG 758
Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDCVCTS 774

RESULT 13
US-10-121-060-142
Sequence 142, Application US/10121060
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C21
CURRENT APPLICATION NUMBER: US/10/121,060
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 142
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-060-142
```

```
Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
Qy 75 CKEFEYPCBEGRCQPVAEQAQESCYRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYIYP--EGECCPV-----CEDPVYFPNPNAGCYANGLILAHGRWRDEDDCT--FCQ 425
```

Qy	117	SVPEYATVGSYPPIELIALGKKCCDVVITQ-----QLPCEAEVSSDP-----ETPTSD	167
Db	426	CYN-----GERHCVAIVCGQTCTNPVKVPGCECPVCEETPIITVDPAC	469
Qy	168	GKLV-----WKIDRLGAGOKCKITVW-----KPLKBCG---C-----FTAAT	202
Db	470	GELSNCTLTGKDINGFKRHNG-----CRTCCQINTEELCSEKQGCFLNCPGFLTDAQ	525
Qy	203	VCACPELASYKCGOPATCIK-----OEGPD-CACLBCP---VCYKIEVNTGSAIA	250
Db	526	NCEICECRPRPKRPIICDKYCPGLLLKNKHGCDICRCKKPELUSCKI-----	575
Qy	251	RNVYDNPVDPGYSHASQORVLSFNGLDMRPGKVFTEFCPPRRGOITNVAIVCYGG	310
Db	576	-----CPGFOQDSHGCLICKRE	594
Qy	311	HKCSANVTITVNEPCVOV-----NISGADMSYVCKPV-----EYSISVSNP-----GDVLV	356
Db	595	ASASAG-PPILSGTCLTVGDGHHKHNEASHHDGCRECYCLNGREMCALITCPVPACNGPTI	653
Qy	357	H-----DVIQTDTLPSGYTVLEAPGCEICCNKVWRI-----KE	390
Db	654	HPQCQPCSADDFVQRPESLTPSICHAPGGEYFVEGETWINDSCTQCTCHSGVLCETE	713
Qy	391	MCPGETLQKLVVAQVPGRETNQAVTSESNGCTCSCAETTHWKGLAATHMVCVLDTN	450
Db	714	VCP-----PLUCQNSR-----TQDSCCPQCTQPPRPSLSRNNSPVNYCKNDEG	758
Qy	451	DPICVGENTVVRIQVT	466
Db	759	DIFLAAESKMPDVCTS	774

RESULT 14

RESOL 14  
US-10-121-063-142  
; Sequence 142, Application US/10121063

1	GENERAL INFORMATION:
2	APPLICANT: Baker, Kevin P.
3	APPLICANT: Barasini, Maureen
4	APPLICANT: DeForge, Laura
5	APPLICANT: Desnoyers, Luc
6	APPLICANT: Flivaroff, Ellen
7	APPLICANT: Gao, Wei-Qiang
8	APPLICANT: Gerritsen, Mary E.
9	APPLICANT: Goddard, Audrey
10	APPLICANT: Godowski, Paul J.
11	APPLICANT: Gurney, Austin L.
12	APPLICANT: Sherwood, Steven
13	APPLICANT: Smith, Victoria
14	APPLICANT: Stewart, Timothy A.
15	APPLICANT: Tumas, Daniel
16	APPLICANT: Watanabe, Colin K
17	APPLICANT: Wood, William

APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;
; FILE REFERENCE: P3330R1C19
;
; CURRENT APPLICATION NUMBER: US/10/121.063
;
; CURRENT FILING DATE: 2002-04-12
;
; Prior Application removed - See file wrapper or Palm
;

```

; Prior Application removed - See File Wrapper or Palm

```

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-063-142

Query Match          3.9%;   Score 114.5;   DB 6;   Length 1036;
Best Local Similarity 19.2%;   Pred. No. 0.75;
Matches 95;   Conservative 45;   Mismatches 155;   Indels 201;   Gaps 24;

QY      75  CDKEFYPCPEGRCQVPEAQESCGRLYSVK-----VNDDCNVEICO 116

```

Db	376	CERYYP---EGECCPV-		425
Qy	117	SVPEYATGSPYPIELAIKKKDCVVIITQ-----QLPCEAEFVSSDP---ETTPSTD	167	
Db	426	CVN-----GERHCVAVCVGQTCTNPVKVPGCCPCVCEPTIITVDPAC	469	
Qy	168	GKLV-----WKIDRLGAGDKCKITVW-----KPLKGC---C---PTAAT	202	
Db	470	GELSNCTLTGKDCINGFKRDHNG-----CRTCCINTEELCSERKOGCTLNCPPGFLTDAQ	525	
Qy	203	VCACPRLSYTKCGOPALCIK-----QEGPD-CACLRCP--VCYKIEVWNTGSAIA	250	
Db	526	NCEICECRPRPKRPIIDCKYCPGLGLLNKHCDCICRCKCPELSSCKI-----	575	
Qy	251	RNVTVDNVPDGYSHASQORVLSNLDMPKDKVFTVEFCQPRRGITNVAITVTCGG	310	
Db	576	-----CPLGFQDSSHGCLICKCRE	594	
Qy	311	HKCSANVTTVNNEPCVQV-----NISGADWSYCKPV-----EYSISVSNP-----GDVLVL	356	
Db	595	ASASAG-PPLISGTCCLTVGDGHHHKNESHWDCRECYCLUNGREMCALITCPVACGNPTI	653	
Qy	357	H-----DVIYQDTLPSPGVTVLEAPGETCCNKVWRI-----KE	390	
Db	654	HPGQCPCPADDFVQKPELSTPISCHAPGEYFVEGETWNIIDSCOTCTCHSGHVLCETE	713	
Qy	391	MCPGETLQKLVYKAQVPGRETNQAVTSESNGTCSCAETTHHWKGLAATHMCVLDTN	450	
Db	714	VCP-----PLLCONPSR-----TODSCCPQCTDQPRPSLSRNSVNYCKNDEG	758	
Qy	451	DPICVGENTVYRIQVT	466	
Db	759	DIFLAAESWKPDVCTS	774	

RESULT 15

US-10-123-108-142  
; Sequence 142, Application US/10123108

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Flavio, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William

APPLICANT: Zhang, Yemin

7. APPLICANT: Zhang, Zhenlin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME





```
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match          3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75; 155; Indels 201; Gaps 24;
Matches 95; Conservative 45; Mismatches 45;

QY 75 CDKEFYPCERGROPEAQEQSCYGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVP--EGECCPV-----CEDPVFPFNNPAGCYANGLILAHGDRNREDDCT--FCQ 425

QY 117 SVPEYATVSPYPIETILAIGKDKCDVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGGTCTNPVKVPGCECPVCEPTIITVDPPAC 469

QY 168 GKLV-----WKIDRLGAGDKCKITVW-----KPLKEGC--C---FTAAT 202
Db 470 GELSNCLTGTCKDINGFKRDHNG---CRTQCINTEELCSERKQGCTLNCPEGFLTDAQ 525

QY 203 VCACPELRSYTKGQPAICIK-----OEGPD-CACLRCP--VCYKIEVVNTGSATA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLKKNKHGCDICRCKKPELSCSKI-----575

QY 251 RNVTVDPVDPGYSHASGQRLVSNLGNMRPGDKKVFTEFCPORRGQITNVATVYCGG 310
Db 576 -----CPLGFOODSHGCLICKCRE 594

QY 311 HKCSANVTYVNEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHHDGCRECYCLNGREMCALITCPVPACGNPTI 653

QY 357 H-----DVVIQDTLPSTVLEAPGGEICGNKVWRI-----KE 390
Db 654 HPGCCPCSCADDFVQKPELSTPSICHAPGGEYFVEGETWNIDSCOTCTCHSGRVLCETE 713

QY 391 MCPGETLQKLVVKAQVPGREFINQAVTSESNCGTCTCAETTHWKGLAATHMVCVLDIN 450
Db 714 VCP-----PLLCQNPSR-----TQDSCCPOCTQDPPRPSLRNNSVNPYCKNDEG 758

QY 451 DPICVGENTVYRICVT 466
Db 759 DIFLAESWKPVDVCTS 774
```

Search completed: May 25, 2002, 22:22:16  
Job time: 463 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:21:04 ; Search time 35.96 Seconds  
(without alignments)  
1717.381 Million cell updates/sec

Title: US-09-523-647-2

Perfect score: 556

Sequence: 1 MSKLIRRVTVTLAITSMA...ILSSDTLTSPVSDTENTHYV 556

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 12

Total number of hits satisfying chosen parameters: 24

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

length of match

Result No.	Score	Match	Length	ID	Description
1	556	100.0	556	20	AA1980.DAT
2	556	100.0	556	21	AA1981.DAT
3	25	4.5	553	20	AA197788
4	16	2.9	16	20	AA192728
5	16	2.9	16	21	AA193677
6	16	2.9	16	21	AA193817
7	16	2.9	16	22	AA193245
8	16	2.9	20	21	AA193674
9	16	2.9	20	22	AA193244
10	13	2.3	109	11	AA197399
11	12	2.2	12	21	AA193676

## ALIGNMENTS

## RESULT 1

AA193676 standard; Protein; 556 AA.

AC AA193676;

DT 13-SEP-1999 (first entry)

DE Chlamydia pneumoniae transmembrane protein sequence.

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.

OS Chlamydia pneumoniae.

PN WO9927105-A2.

PD 03-JUN-1999.

PF 20-NOV-1998; 98WO-1B01890.

PR 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

PA (GEST) GENSET.

PI Griffais R;

WPI; 1999-357842/30.

PT Genome sequence of Chlamydia pneumoniae

PS Page 1037-1038; Disclosure; 1912pp; English.

CC AA193676 represent the proteins encoded by all the open reading  
frames in the complete genome (see AA193676) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
frames of the C. pneumoniae genome (see AA193676) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.

SQ Sequence 556 AA;

Query Match 100.0%; Score 556; DB 20; Length 556;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVTLAITSMA...ILSSDTLTSPVSDTENTHYV 556

12 12 2.2 12 22 AAG83244 Chlamydia trachoma  
13 12 2.2 16 21 AAB13675 C. trachomatis OMC  
14 12 2.2 16 22 AAG83243 Chlamydia trachoma  
15 12 2.2 20 21 AAB13665 C. trachomatis OMC  
16 12 2.2 20 21 AAB13666 C. trachomatis OMC  
17 12 2.2 20 21 AAB13670 C. trachomatis OMC  
18 12 2.2 20 21 AAB13671 C. trachomatis OMC  
19 12 2.2 20 21 AAB13673 Chlamydia trachoma  
20 12 2.2 20 22 AAG83233 Chla  
21 12 2.2 20 22 AAG83234 Chla  
22 12 2.2 20 22 AAG83238 Chla  
23 12 2.2 20 22 AAG83239 Chla  
24 12 2.2 20 22 AAG83241 Chla

```

Db 1 msklrrvvtvlatmsacsfasggieaavaeslltklvasaetkppapmtakkrvlvr 60
Qy 61 RNKQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120
Db 61 rnkqpvdksgatcdkefyppceegrqcpveaqqescygrlysvkvnddcnveicqsvpe 120
Qy 121 YATVGSPPYPIETLAIKGDVVDVITQOLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
Db 121 yatvgsppyeiellaigkdkcdvvtqglpceaeefvssdpetttsdgklwklidrlgag 180
Qy 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGOPAIKJQSGPDCACLRCPVCYKI 240
Db 181 dkckitvwwkplkegcgcftaatvcacpelrsytkcgopaicjkqsgpdcacrlrcpvcyki 240
Qy 241 EVVNTGSATARNVTVDNVPDGYSHASQRLVLSFNLGDMRPGDKKVFVFCPORRGQIT 300
Db 241 evvntgsalarntvtdnvpogyshasqgrvlisfnlgdmrpgdkkvvfvefcporrgqit 300
Qy 301 NVATVYCGGHKCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 301 nvatvycgghkcsanvttvvnepcvovnisgadvsvyckpveyisvsnpgdlvlhdvv 360
Qy 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVYKQVPCRFNTQAVTSE 420
Db 361 iqdtlpsgvttvleapggeiccnkvvrkempcgetlqflvkaqvpgfrntqavtse 420
Qy 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTRNGSAEDTNVSLIL 480
Db 421 sngctctcaetthwkglaathmclvldndpicvgentvyrilcvtrngsaedtnvslil 480
Qy 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKEVSFEVSTLKGIAPGDARGEAILSS 540
Db 481 kfskelqpiassgptktisgntvvfdalpklgskesvefsvtlkgiapgdargeailss 540
Qy 541 DTLTSPVSDTENTHVV 556
Db 541 dtltspvsdtenhv 556

RESULT 2
AAB18820
ID AAB18820 standard; Protein; 556 AA.
AC AAB18820;
DT 22-JAN-2001 (first entry)
DE A 60 kda cysteine-rich membrane protein of Chlamydia pneumoniae.
KW Cysteine-rich membrane protein; Chlamydia infection; bronchitis;
KW community acquired pneumonia; upper respiratory tract infection; vaccine;
KW sinusitis.
OS Chlamydia pneumoniae.
XX WO200055326-A1.
PN 21-SEP-2000.
XX 09-MAR-2000; 2000WO-CA00240.
XX 12-MAR-1999; 99US-0123966.
XX (AVET ) AVENTIS PASTEUR LTD.
XX Murdin AD, Oomen RP, Wang J, Dunn P;
XX WPI; 2000-618918/59.
XX N-PSDB; AAB75901.
XX New polynucleotides encoding a 60kda cysteine-rich membrane protein
XX from Chlamydia, useful as a vaccine for preventing and treating
PT

```

```

Pt Chlamydia infection in mammals -
XX Claim 16; Fig 1; 77pp; English.
XX The present sequence represents a Chlamydia 60 kda cysteine-rich membrane
CC protein. The membrane-rich polynucleotide and polypeptide are useful
CC for preventing or treating Chlamydia infection, such as community
CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
CC They are also useful for diagnosing Chlamydia infection by assaying a
CC body fluid of a mammal. The polypeptide is useful for vaccine
CC production.
XX Sequence 556 AA;
SQ

Query Match 100.0%; Score 556; DB 21; Length 556;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLIRRVVTVLALTSMASCFASGGIEAAVAESLITKIVASAETKPAPVPMATAKRVLR 60
Db 1 msklirrvvtvlatmsacsfasggieaavaeslltklvasaetkppapmtakkrvlvr 60
Qy 61 RNKQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120
Db 61 rnkqpvdksgatcdkefyppceegrqcpveaqqescygrlysvkvnddcnveicqsvpe 120
Qy 121 YATVGSPPYPIETLAIKGDVVDVITQOLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
Db 121 yatvgsppyeiellaigkdkcdvvtqglpceaeefvssdpetttsdgklwklidrlgag 180
Qy 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGOPAIKJQSGPDCACLRCPVCYKI 240
Db 181 dkckitvwwkplkegcgcftaatvcacpelrsytkcgopaicjkqsgpdcacrlrcpvcyki 240
Qy 241 EVVNTGSATARNVTVDNVPDGYSHASQRLVLSFNLGDMRPGDKKVFVFCPORRGQIT 300
Db 241 evvntgsalarntvtdnvpogyshasqgrvlisfnlgdmrpgdkkvvfvefcporrgqit 300
Qy 301 NVATVYCGGHKCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 301 nvatvycgghkcsanvttvvnepcvovnisgadvsvyckpveyisvsnpgdlvlhdvv 360
Qy 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVYKQVPCRFNTQAVTSE 420
Db 361 iqdtlpsgvttvleapggeiccnkvvrkempcgetlqflvkaqvpgfrntqavtse 420
Qy 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTRNGSAEDTNVSLIL 480
Db 421 sngctctcaetthwkglaathmclvldndpicvgentvyrilcvtrngsaedtnvslil 480
Qy 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKEVSFEVSTLKGIAPGDARGEAILSS 540
Db 481 kfskelqpiassgptktisgntvvfdalpklgskesvefsvtlkgiapgdargeailss 540
Qy 541 DTLTSPVSDTENTHVV 556
Db 541 dtltspvsdtenhv 556

RESULT 3
AAY37788
ID AAY37788 standard; Protein; 553 AA.
XX AAY37788;
XX AAY37788;
XX 07-OCT-1999 (first entry)
XX Chlamydia trachomatis cellular envelope protein.
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
XX paratrachoma; inclusion conjunctivitis; genital disease; perilihepatitis;
XX nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW

```

KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

OS Chlamydia trachomatis.

XX WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-TB01939.

XX 04-NOV-1998; 98US-0107077.

XX 28-NOV-1997; 97FR-0015041.

XX 17-DEC-1997; 97FR-0016034.

XX (GEST ) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure; Page 1371-1372; 1755pp; English.

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
can also be used to control growth of the microorganism. Chlamydia  
trachomatis is responsible for a large number of diseases, e.g. eye  
diseases such as conventional trachoma, nonendemic trachoma,  
paratrachoma, and inclusion conjunctivitis; genital diseases such as  
nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
depression, bartholinitis; pneumopathy in breast feeding infants;  
and venereal lymphogranulomatosis. The polypeptides of the invention  
may be of use in treating these diseases.

XX Sequence 553 AA;

Query Match 4.5%; Score 25; DB 20; Length 553;

Best Local Similarity 100.0%; Pred. No. 9.4e-17;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWKPKEGCGCFTAAATVCACPE 208

Db 181 kitvwwkpkegcgcftaatvcacpe 205

RESULT 4

AAI42728

ID AAI42728 standard; peptide; 16 AA.

XX AC AAY42728;

XX 20-DEC-1999 (first entry)

XX Chlamydia pneumoniae outer membrane protein-derived peptide.

XX Heart disease; inflammatory; autoimmune; cardiomyopathy; myosin;

XX Chlamydia; induction; vaccine.

XX Synthetic.

XX Chlamydia pneumoniae.

XX Key Location/Qualifiers

FT Modified-site 1

FT Misc-difference 16 /note= "optionally N-terminally acetylated"

XX US5962636-A.

XX 05-OCT-1999.

XX 12-AUG-1998; 98US-0133774.

XX 12-AUG-1998; 98US-0133774.

XX (AMGE-) AMGEN CANADA INC.

XX Bachmaier K, Hessel AJ, Penninger JM, Neu N;

XX WPI; 1999-589735/50.

XX N-PSDB; AAZ28184.

XX Peptides that induce or suppress inflammatory cardiomyopathy

XX Claim 1; Column 19; 17pp; English.

XX This sequence represents Chlamydia pneumoniae outer membrane protein-  
derived peptide, which induces inflammatory cardiomyopathy (ICM) via  
an autoimmune response in mice immunised with it. It contains an amino  
acid sequence motif MaxxxS (AAI42722) which appears to be required for  
the induction of this disease. The motif was originally identified in  
the sequence of murine alpha myosin heavy chain-derived peptide,  
M7A-alpha, when it was compared with a peptide derived from a homologous  
region of the murine beta myosin heavy chain, M7A-beta (AAI42724) which  
did not cause the disease. Several peptide fragments containing the  
motif were identified from a database and were found to be fragments of  
cysteine rich outer membrane proteins from various species of Chlamydia.  
These peptides also induced ICM, indicating that infection with  
Chlamydia may be involved in the development of ICM. Inflammatory  
cardiomyopathy peptides are used to determine the risk of ICM by  
incubation with a subject's T cells and measuring the degree of  
proliferation (an increased degree being indicative of risk) or to raise  
specific antibodies which can be used therapeutically and for the  
detection of Chlamydia. Such peptides can also be used with an adjuvant  
and an excipient in a vaccine for decreasing ICM.

XX Sequence 16 AA;

Query Match 2.9%; Score 16; DB 20; Length 16;

Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GIEAAVAESLITKIVA 40

Db 1 gieaavaeslitriva 16

RESULT 5

AAI3677

ID AAI3677 standard; Peptide; 16 AA.

XX AC AAI3677;

XX 02-FEB-2001 (first entry)

XX C. pneumoniae OMCB peptide 185-198.

XX Chlamydial infection; sexually transmitted disease;

XX pelvic inflammatory disease; PID; tubal obstruction; infertility;

XX trachoma; blindness; acute respiratory tract infection;

XX atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia pneumoniae.

XX WO200034483-A2.

XX 15-JUN-2000.

XX 08-DEC-1999; 99WO-US29012.

XX 08-DEC-1998; 98US-0208277.

XX 08-APR-1999; 99US-0288594.

PR 01-OCT-1999; 99US-0410568.  
 PR 22-OCT-1999; 99US-0426571.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
 XX  
 DR WPI: 2000-431303/37.  
 XX  
 XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises  
 PT amino acid sequence encoded by polynucleotide sequence -  
 XX  
 XX Disclosure; Page 233; 256pp; English.  
 PS  
 XX The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydiae are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present invention is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 2.9%; Score 16; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 180 GDKCKIIVWVKPLKEG 195  
 DB 1 gdkckitvwvklkeg 16  
 |||||  
 RESULT 6  
 AAY83817  
 ID AAY83817 standard; Protein; 16 AA.  
 XX  
 AC AAY83817;  
 XX  
 DT 21-JUN-2000 (first entry)  
 XX  
 DE Chlamydia pneumoniae Cys-rich OMP peptide.  
 XX  
 KW Cardiant; murine alpha myosin heavy chain; inflammatory myocarditis;  
 KW autoimmune inflammatory cardiomyopathy; Chlamydia; antibody; vaccine;  
 KW hybridization probe.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 PN US6034230-A.  
 XX  
 PD 07-MAR-2000.  
 XX  
 PF 03-MAY-1999; 99US-0303862.  
 XX  
 PR 12-AUG-1998; 98US-0133774.  
 XX  
 XX (AMGE-) AMGEN CANADA INC.  
 PA  
 XX Neu N, Penninger JM, Bachmaier K, Hessel AJ;  
 PI  
 XX WPI: 2000-255712/22.  
 DR  
 DR N-PSDB: AAZ99167.  
 XX

PT DNA molecules encoding novel myocardial peptides used for inhibiting  
 PT and inducing inflammatory cardiomyopathy in vivo -  
 XX  
 XX Example 1; Column 18; 17pp; English.  
 XX  
 CC This sequence represents a fragment of the Chlamydia pneumoniae Cys-rich  
 CC outer membrane protein (OMP). The sequence was obtained by homology  
 CC searches of the PIR database using the murine M7A-alpha/beta peptide  
 CC consensus sequence (Y83813) as the query. The screen isolated the  
 CC peptides Y83814-Y83819 and their corresponding coding sequences  
 CC 259164-259169. The invention relates to the isolation of sequences  
 CC coding for peptide sequences derived from bacteria and viruses which may  
 CC cause inflammatory cardiomyopathy. The peptide sequences are searched  
 CC based on the sequence of the M7a peptides derived from the murine alpha  
 CC myosin heavy chain polypeptide. The peptides encoded by the DNAs are  
 CC used, alone or in conjunction with other therapeutics, for inducing or  
 CC inhibiting inflammatory cardiomyopathy in vivo, where the cardiomyopathy  
 CC is autoimmune inflammatory cardiomyopathy, and inflammatory  
 CC cardiomyopathy caused by Chlamydia or other bacterial or viral infections  
 CC that cause inflammatory myocarditis in a mammal. Antibodies against the  
 CC peptides and the peptides themselves are used for measuring the risk of  
 CC inflammatory cardiomyopathy in a mammal. The peptides may also be used  
 CC in vaccines. Nucleic acids encoding the peptides may be used as  
 CC hybridization probes, e.g. in diagnostic assays to test for the  
 CC presence of Chlamydia DNA.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 2.9%; Score 16; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 25 GIEAAVAESLITKIYA 40  
 DB 1 gieaavaeslilkiya 16  
 |||||  
 RESULT 7  
 AAG83245  
 ID AAG83245 standard; Peptide; 16 AA.  
 XX  
 AC AAG83245;  
 XX  
 DT 05-SEP-2001 (first entry)  
 XX  
 DE Chlamydia trachomatis OMCB peptide 185-198.  
 XX  
 KW Chlamydia; vaccine; infection; fusion protein; antigen;  
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
 KW acute respiratory tract infection; Capi; C7529; OMCB;  
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 PN WC200140474-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 04-DEC-2000; 2000WO-US32919.  
 XX  
 PR 03-DEC-1999; 99US-0454684.  
 PR 19-APR-2000; 2000US-0556877.  
 PR 20-JUN-2000; 2000US-0598419.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;  
 PI  
 XX WPI: 2001-374831/39.  
 DR  
 XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic

PT inflammatory disease, trachoma, acute respiratory tract infections,  
PT atherosclerosis and heart disease -  
PS Claim 66; Page 239; 295pp; English.  
XX  
CC The present peptide is provided in a specification relating to  
CC compounds and methods for the treatment and diagnosis of chlamydial  
CC infection. The compounds provided include polypeptides and fusion  
CC proteins comprising immunogenic portions of Chlamydia antigens  
CC and DNA sequences encoding such polypeptides. They are useful for  
CC vaccinating against chlamydial infection, which causes pelvic  
CC inflammatory disease, trachoma, acute respiratory tract infections,  
CC atherosclerosis and heart disease.  
XX  
SQ Sequence 16 AA;  
  
Query Match 2.9%; Score 16; DB 22; Length 16;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0;  
  
QY 180 GDKCKITVWVKPLKEG 195  
DB 1 gdkckitvwvkplkeg 16  
|||||  
  
RESULT 8  
AAB13674  
ID AAB13674 standard; Peptide; 20 AA.  
XX  
AC AAB13674;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE C. trachomatis OMCB peptide 171-190.  
XX  
KW Chlamydial infection; sexually transmitted disease;  
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
KW trachoma; blindness; acute respiratory tract infection;  
KW atherosclerosis; coronary heart disease; antibacterial.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO200034483-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 08-DEC-1999; 99WO-US29012.  
XX  
PR 08-DEC-1998; 98US-0208277.  
PR 08-APR-1999; 99US-0288594.  
PR 01-OCT-1999; 99US-0410568.  
PR 22-OCT-1999; 99US-0426571.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
XX  
XX WPI; 2000-431303/37.  
XX  
XX  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
XX comprises immunogenic portion of Chlamydia antigen, which comprises  
XX amino acid sequence encoded by polynucleotide sequence -  
XX  
XX Disclosure; Page 233; 256pp; English.  
XX  
XX The present invention relates to new nucleic acid sequences and the  
XX proteins encoded by the nucleic acid sequences. The encoded proteins  
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded  
XX proteins are useful for the serodiagnosis and treatment of Chlamydia  
XX infection. Chlamydiae are intracellular bacterial pathogens that are  
XX responsible for a wide variety of human infections. C. trachomatis  
XX infection is one of the most common sexually transmitted diseases and can

CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present sequence is a protein isolated in the  
CC present invention.  
XX  
SQ Sequence 20 AA;  
  
Query Match 2.9%; Score 16; DB 21; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1e-08; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0;  
  
QY 184 KITVWVKPLKEGCGFT 199  
DB 5 kitvwvkplkegcft 20  
|||||  
  
RESULT 9  
AAG83242  
ID AAG83242 standard; Peptide; 20 AA.  
XX  
AC AAG83242;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Chlamydia trachomatis OMCB peptide 171-190.  
XX  
KW Chlamydia; vaccine; infection; fusion protein; antigen;  
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
KW acute respiratory tract infection; Capl; CF529; OMCB; thiol specific antioxidant; TSA.  
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO200140474-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 04-DEC-2000; 2000WO-US32919.  
XX  
PR 03-DEC-1999; 99US-0454684.  
PR 19-APR-2000; 2000US-0556877.  
PR 20-JUN-2000; 2000US-0598419.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;  
XX  
XX WPI; 2001-374831/39.  
XX  
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
XX inflammatory disease, trachoma, acute respiratory tract infections,  
XX atherosclerosis and heart disease -  
XX  
XX Claim 66; Page 239; 295pp; English.  
XX  
XX The present peptide is provided in a specification relating to  
XX compounds and methods for the treatment and diagnosis of chlamydial  
XX infection. The compounds provided include polypeptides and fusion  
XX proteins comprising immunogenic portions of Chlamydia antigens  
XX and DNA sequences encoding such polypeptides. They are useful for  
XX vaccinating against chlamydial infection, which causes pelvic  
XX inflammatory disease, trachoma, acute respiratory tract infections,  
XX atherosclerosis and heart disease.  
XX  
SQ Sequence 20 AA;  
  
Query Match 2.9%; Score 16; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1e-08; Mismatches 0; Indels 0; Gaps 0;  
Matches 16; Conservative 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVRLKGGCFT 199  
Db 5 kitvwwkplkegcft 20

## RESULT 10

AAR07399  
ID AAR07399 standard; protein; 109 AA.

AC AAR07399;

DT 12-MAR-1991 (first entry)

DE Chlamydia Crp gene prod.

KW PCR; DIANA; solid phase medical diagnosis.

OS Chlamydia trachomatis.

XX WO9011369-A.

PD 04-OCT-1990.

PF 15-MAR-1990; 90WO-EP00454.

PR 22-MAR-1989; 89GB-0006642.

PR 22-MAR-1989; 89GB-0006641.

PA (CEMU-) CEMU BIOTEKNIK.

XX Uhlen M;

DR WPI; 1990-320269/42.

DR N-FSDB; AAQ06236.

XX Amplification of DNA by PCR - using distal DNA on a primer as a

PT handle for attachment to a solid support or a label, used in

PT diagnosis of medical conditions

XX Disclosure; fig 19; 66pp; English.

XX This is the product encoded by the Crp gene. The corresp. gene can

CC be detected utilising DIANA, an improved PCR method with 2 primer

CC pairs.

XX Sequence 109 AA;

QY 332 GADWSYVCKPVEY 344

Db 30 gadwsyvcckpvey 42

## RESULT 11

AAB13676  
ID AAB13676 standard; Peptide; 12 AA.

AC AAB13676;

DT 02-FEB-2001 (first entry)

DE C. trachomatis OMCB peptide 175-186.

XX Chlamydial infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;

KW trachoma; blindness; acute respiratory tract infection;

KW atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia trachomatis.  
OS WO200034483-A2.  
PN 15-JUN-2000.  
XX

PD

XX 08-DEC-1999; 99WO-US29012.

PF 08-DEC-1999; 98US-0208277.

PR 08-APR-1999; 99US-0288594.

PR 01-OCT-1999; 99US-0410568.

PR 22-OCT-1999; 99US-0426571.

XX (CORI-) CORIXA CORP.

XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX WPI; 2000-431303/37.

DR

XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection

XX comprises immunogenic portion of Chlamydia antigen, which comprises

XX amino acid sequence encoded by polynucleotide sequence .

XX Disclosure; Page 233; 256pp; English.

XX The present invention relates to new nucleic acid sequences and the

CC proteins encoded by the nucleic acid sequences. The encoded proteins

CC comprise an immunogenic portion of a Chlamydia antigen. The encoded

CC proteins are useful for the serodiagnosis and treatment of Chlamydia

CC infection. Chlamidia are intracellular bacterial pathogens that are

CC responsible for a wide variety of human infections. C. trachomatis

CC infection is one of the most common sexually transmitted diseases and can

CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction

CC and infertility. Trachoma due to ocular infection with C. trachomatis is

CC the leading cause of preventable blindness worldwide. C. pneumonia is a

CC major cause of acute respiratory tract infections in humans and is also

CC thought to play a role in the pathogenesis of atherosclerosis and

CC coronary heart disease. The present sequence is a protein isolated in the

CC present invention.

XX Sequence 12 AA;

SQ

Query Match 2.2%; Score 12; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 8.5e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVRLKKEG 195

Db 1 kitvwwkplkeg 12

## RESULT 12

AAG83244

ID AAG83244 standard; Peptide; 12 AA.

XX AAG83244;

XX 05-SEP-2001 (first entry)

XX Chlamydia trachomatis OMCB peptide 175-186.

XX Chlamydia; vaccine; infection; fusion protein; antigen;

KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;

KW acute respiratory tract infection; CapJ; CT529; OMCB;

KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

XX Chlamydia trachomatis.

XX WO200140474-A2.

XX 07-JUN-2001.



```
XX PF 04-DEC-2000; 2000WO-US32919.
XX PR 03-DEC-1999; 99US-0454684.
XX PR 19-APR-2000; 2000US-0556877.
XX PR 20-JUN-2000; 2000US-0598419.
XX PA (CORI-) CORIXA CORP.
XX PJ Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX DF WPI; 2001-374831/39.
XX PA Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX PR inflammatory disease, trachoma, acute respiratory tract infections,
XX PR atherosclerosis and heart disease -
XX PF Claim 66; Page 239; 295pp; English.
XX CC The present peptide is provided in a specification relating to
XX CC compounds and methods for the treatment and diagnosis of chlamydial
XX CC infection. The compounds provided include polypeptides and fusion
XX CC proteins comprising immunogenic portions of Chlamydia antigens
XX CC and DNA sequences encoding such polypeptides. They are useful for
XX CC vaccinating against chlamydial infection, which causes pelvic
XX CC inflammatory disease, trachoma, acute respiratory tract infections,
XX CC atherosclerosis and heart disease.
XX SQ Sequence 12 AA;
XX QY Query Match 2.2%; Score 12; DB 22; Length 12;
XX DB Best Local Similarity 100.0%; Pred. No. 8.5e-05;
XX M Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 184 KITVWVKPLKEG 195
XX DB | | | | | | | | | |
XX AC 1 kitvwvkplkeg 12
XX DT 02-FEB-2001 (first entry)
XX DE C. trachomatis OMCB peptide 171-186.
XX KW Chlamydial infection; sexually transmitted disease;
XX KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX KW trachoma; blindness; acute respiratory tract infection;
XX KW atherosclerosis; coronary heart disease; antibacterial.
XX OS Chlamydia trachomatis.
XX PN WO200034483-A2.
XX XX 15-JUN-2000.
XX PF 08-DEC-1999; 99WO-US29012.
XX PR 08-DEC-1998; 98US-0208277.
XX PR 08-APR-1999; 99US-0288594.
XX PR 01-OCT-1999; 99US-0410568.
XX PR 22-OCT-1999; 99US-0426571.
XX PA (CORI-) CORIXA CORP.
XX PJ Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX PR WPI; 2000-431303/37.
XX DR

XX PF Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX PR comprises immunogenic portion of Chlamydia antigen, which comprises
XX PT amino acid sequence encoded by polynucleotide sequence -
XX XX Disclosure; Page 233; 256pp; English.
XX CC The present invention relates to new nucleic acid sequences and the
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia
XX CC infection. Chlamydiae are intracellular bacterial pathogens that are
XX CC responsible for a wide variety of human infections. C. trachomatis
XX CC infection is one of the most common sexually transmitted diseases and can
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is
XX CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
XX CC major cause of acute respiratory tract infections in humans and is also
XX CC thought to play a role in the pathogenesis of atherosclerosis and
XX CC coronary heart disease. The present sequence is a protein isolated in the
XX CC present invention.
XX SQ Sequence 16 AA;
XX QY Query Match 2.2%; Score 12; DB 21; Length 16;
XX DB Best Local Similarity 100.0%; Pred. No. 0.00011;
XX M Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 184 KITVWVKPLKEG 195
XX DB | | | | | | | | | |
XX AC 5 kitvwvkplkeg 16
XX DT 05-SEP-2001 (first entry)
XX DE Chlamydia trachomatis OMCB peptide 171-186.
XX KW Chlamydia; vaccine; infection; fusion protein; antigen;
XX KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
XX KW acute respiratory tract infection; Capi; CF529; OMCB;
XX KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX OS Chlamydia trachomatis.
XX PN WO200140474-A2.
XX XX 07-JUN-2001.
XX PF 04-DEC-2000; 2000WO-US32919.
XX PR 03-DEC-1999; 99US-0454684.
XX PR 19-APR-2000; 2000US-0556877.
XX PR 20-JUN-2000; 2000US-0598419.
XX PA (CORI-) CORIXA CORP.
XX PJ Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX XX WPI; 2001-374831/39.
XX PR Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX PR inflammatory disease, trachoma, acute respiratory tract infections,
XX PR atherosclerosis and heart disease -
XX XX Claim 66; Page 239; 295pp; English.
XX DR
```

CC The present peptide is provided in a specification relating to  
CC compounds and methods for the treatment and diagnosis of chlamydial  
CC infection. The compounds provided include polypeptides and fusion  
CC proteins comprising immunogenic portions of Chlamydia antigens  
CC and DNA sequences encoding such polypeptides. They are useful for  
CC vaccinating against chlamydial infection, which causes pelvic  
CC inflammatory disease, trachoma, acute respiratory tract infections,  
CC atherosclerosis and heart disease.

XX Sequence 16 AA;

Query Match 2.2%; Score 12; DB 22; Length 16;

Best Local Similarity 100.0%; Pred. No. 0.00011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEG 195

DB 5 KITVWVKPLKEG 16

#### RESULT 15

AAB13665  
ID AAB13665 standard; Peptide: 20 AA.

XX AC AAB13665;

XX DT 02-FEB-2001 (first entry)

XX DE C. trachomatis OMCB peptide 128-147.

XX KW Chlamydial infection; sexually transmitted disease;

XX KW pelvic inflammatory disease; PID; tubal obstruction; infertility;

XX KW trachoma; blindness; acute respiratory tract infection;

XX KW atherosclerosis; coronary heart disease; antibacterial.

XX OS Chlamydia trachomatis.

XX PN WQ200034483-A2.

XX PD 15-JUN-2000.

XX PF 08-DEC-1999; 99WO-US29012.

XX PR 08-DEC-1998; 98US-0208277.

XX PR 08-APR-1999; 99US-0288594.

XX PR 01-OCT-1999; 99US-0410568.

XX PR 22-OCT-1999; 99US-0426571.

XX PA (CORI-) CORIXA CORP.

XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX PS WPI; 2000-431303/37.

XX DR Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
XX PT comprises immunogenic portion of Chlamydia antigen, which comprises  
XX PT amino acid sequence encoded by polynucleotide sequence -

XX PS Disclosure; Page 230; 256pp; English.

XX CC The present invention relates to new nucleic acid sequences and the  
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins  
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
XX CC infection. Chlamydiae are intracellular bacterial pathogens that are  
XX CC responsible for a wide variety of human infections. C. trachomatis  
XX CC infection is one of the most common sexually transmitted diseases and can  
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
XX CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
XX CC major cause of acute respiratory tract infections in humans and is also  
XX CC thought to play a role in the pathogenesis of atherosclerosis and

CC coronary heart disease. The present sequence is a protein isolated in the  
CC present invention.

XX Sequence 20 AA;

Query Match 2.2%; Score 12; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ITQQLPCEAEFV 156

DB 9 ITQQLPCEAEFV 20

Search completed: May 25, 2002, 22:25:32

Job time: 268 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:22:22 ; Search time 16.39 Seconds  
(without alignments)  
828.592 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 556  
Sequence: 1 MSKLRRVTVLALRSMASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 231628 seqs, 2442594 residues

Word size: 12

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	4.5	25	4 US-09-025-596-38	Sequence 38, Appl
2	20	3.6	20	4 US-09-025-596-41	Sequence 41, Appl
3	19	3.4	19	4 US-09-025-596-39	Sequence 39, Appl
4	16	2.9	16	2 US-09-133-774-7	Sequence 7, Appl
5	16	2.9	16	3 US-09-303-862-7	Sequence 7, Appl
6	16	2.9	16	4 US-09-025-596-43	Sequence 43, Appl
7	13	2.3	19	4 US-09-025-596-40	Sequence 40, Appl
8	13	2.3	109	1 US-08-477-270-20	Sequence 20, Appl

#### ALIGNMENTS

RESULT 1  
US-09-025-596-38  
; Sequence 38, Application US/09025596  
; Patent No. 6340463  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, William M.  
; APPLICANT: Stratton, Charles W.  
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE  
; TITLE OF INVENTION: SEQUENCES  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/09/025.596  
; CURRENT FILING DATE: 1998-02-18  
; EARLIER APPLICATION NUMBER: 08/911.593  
; EARLIER FILING DATE: 1997-08-14

; EARLIER APPLICATION NUMBER: 60/023.921  
; EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 38  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-025-596-38

Query Match 4.5%; Score 25; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RNNQPVQKSRGAFCDKEFYPCEE 84  
|||||  
DB 1 RNNQPVQKSRGAFCDKEFYPCEE 25

RESULT 2  
US-09-025-596-41  
; Sequence 41, Application US/09025596  
; Patent No. 6340463  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, William M.  
; APPLICANT: Stratton, Charles W.  
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE  
; TITLE OF INVENTION: SEQUENCES  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/09/025.596  
; CURRENT FILING DATE: 1998-02-18  
; EARLIER APPLICATION NUMBER: 08/911.593  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 60/023.921  
; EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 41  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-025-596-41

Query Match 3.6%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.2e-13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 418 TSNSNCGTCTSCAETTHWK 437  
|||||  
DB 1 TSNSNCGTCTSCAETTHWK 20

RESULT 3  
US-09-025-596-39  
; Sequence 39, Application US/09025596  
; Patent No. 6340463  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, William M.  
; APPLICANT: Stratton, Charles W.  
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE  
; TITLE OF INVENTION: SEQUENCES  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/09/025.596  
; CURRENT FILING DATE: 1998-02-18  
; EARLIER APPLICATION NUMBER: 08/911.593  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 60/023.921  
; EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39

```
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-39

Query Match          3.4%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 278 DMRPGDKKVFVFCPQRR 296
Db 1 DMRPGDKKVFVFCPQRR 19

RESULT 4
US-09-133-774-7
; Sequence 7, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Heart
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-133-774-7

Query Match          2.9%; Score 16; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GIEAAVAESLITKIVA 40
Db 1 GIEAAVAESLITKIVA 16

RESULT 5
US-09-303-862-7
; Sequence 7, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 6034230el Peptides Capable of Modulating Inflammatory Heart
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/303,862
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 09/133,774
; EARLIER FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-303-862-7

Query Match          2.9%; Score 16; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 SSDPETTPTSDCK 169
Db 1 SSDPETTPTSDCK 169

Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GIEAAVAESLITKIVA 40
Db 1 GIEAAVAESLITKIVA 16

RESULT 6
US-09-025-596-43
; Sequence 43, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-43

Query Match          2.9%; Score 16; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 459 TVYRICVTNRGSAEDT 474
Db 1 TVYRICVTNRGSAEDT 16

RESULT 7
US-09-025-596-40
; Sequence 40, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-40

Query Match          2.3%; Score 13; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 SSDPETTPTSDCK 169
Db 1 SSDPETTPTSDCK 169
```

us-09-523-647-2.olig.ra

Tue May 28 08:55:12 2002

Db 1 SSDPETPTSDGK 13

RESULT 8

US-08-477-270-20  
Sequence 20, Application US/08477270  
Patent No. 5629158

GENERAL INFORMATION:

APPLICANT: UHLEN, Mathias

TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL

CONDITIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/477,270

APPLICATION NUMBER: US/08/477,270

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/261,010

FILING DATE:

APPLICATION NUMBER: US 07/781,157

FILING DATE: 07-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 16787/153 DFBC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 109 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-270-20

Query Match 2.3%; Score 13; DB 1; Length 109;

Best Local Similarity 100.0%; Pred. No. 1.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 GADWSYCKPVEY 344

Db 30 GADWSYCKPVEY 42

Search completed: May 25, 2002, 22:25:58

Job time: 216 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 25, 2002, 22:24:02 ; Search time 21.52 seconds  
(without alignments)  
2482.605 Million cell updates/sec

Title: US-09-523-647-2

Parfait score: 556

Sequence: 1 MSKLRRVTVLALTSMA...ILSSDTLSPVSDTENTHYV 556

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Wrd size: 12

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR 71.\*  
1: Pirl.\*  
2: Pirl.\*  
3: Pirl.\*  
4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	100.0	556	1 S12602	60K cysteine-rich
2	556	100.0	556	2 A86560	60 kDa Cysteine-ri
3	45	8.3	557	1 B39439	60K cysteine-rich
4	45	8.1	558	2 JC5204	60K cysteine-rich
5	39	7.0	554	2 C81671	60 kDa outer membr
6	25	4.5	547	1 A32444	60K cysteine-rich
7	25	4.5	547	2 B43584	60K cysteine-rich
8	25	4.5	553	2 D71515	60K cysteine-rich

#### ALIGNMENTS

RESULT 1  
S12602  
60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydia pneumoniae  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 30-Sep-1993 #sequence revision 27-Jun-1994 #text\_change 11-May-2000  
C:Accession: S12602; H72063; A81604  
R:Watson, M.W.; Al-Mandawi, S.; Damden, P.R.; Clarke, I.N.  
Nucleic Acids Res. 19, 5299, 1990  
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of Ch  
A:Reference number: S12602; MUID:90384850  
A:Accession: S12602  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-556 <NAT>  
A:Cross-references: GB:X53511; NID:g550564; PIDN:CAA37590.1; PID:g550566  
A:Experimental source: isolate IOL-207

R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: H72063  
A:Molecule type: DNA  
A:Residues: 1-556 <NAT>  
A:Cross-references: GB:AE001640; GB:AE001363; NID:g4376845; PIDN:AAJ18697.1; PID:g4377  
A:Experimental source: strain CWL029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: A81604  
A:Molecule type: DNA  
A:Residues: 1-556 <NAT>  
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Comment: This protein is associated with the differentiation of reticulate bodies i  
It may also be an important virulence factor.  
C:Genetics:  
A:Gene: omcB; CP0195  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <NAT>

Query Match 100.0% Score 556; DB 1; Length 556;  
Best Local Similarity 100.0%; Pred. No: 0;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLRRVTVLALTSMA...ILSSDTLSPVSDTENTHYV 556  
Db 1 MSKLRRVTVLALTSMA...ILSSDTLSPVSDTENTHYV 556  
Qy 61 RNKQPEQKSRGAFCDKEFYPCGEGRQCPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120  
Db 61 RNKQPEQKSRGAFCDKEFYPCGEGRQCPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120  
Qy 121 YATVGSYPYETILAIKGCDCVDVITQOLPCEAEFVSSDPETPTSDGKLWVXIDRLGAG 180  
Db 121 YATVGSYPYETILAIKGCDCVDVITQOLPCEAEFVSSDPETPTSDGKLWVXIDRLGAG 180  
Qy 181 DKCKITWVKPLKEGCGCFFTAATVCACPELRSYTKCGOPATCIKQEGPDCACLPVCYKI 240  
Db 181 DKCKITWVKPLKEGCGCFFTAATVCACPELRSYTKCGOPATCIKQEGPDCACLPVCYKI 240  
Qy 241 EVNTGSAIARNVTVNDVPDGYSHASGORVLSNLDGMRPDKKVFTEFFCQRGQIT 300  
Db 241 EVNTGSAIARNVTVNDVPDGYSHASGORVLSNLDGMRPDKKVFTEFFCQRGQIT 300  
Qy 301 NVATVTCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGLVLHDVV 360  
Db 301 NVATVTCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGLVLHDVV 360  
Qy 361 IQDTLPSGVTVLEAPGGEICCNKVVNRKEMCPGETIQLQFLVKAQVPGRTTNOVAVTSE 420  
Db 361 IQDTLPSGVTVLEAPGGEICCNKVVNRKEMCPGETIQLQFLVKAQVPGRTTNOVAVTSE 420  
Qy 421 SNGGTCSCAETTHWKGAAATHMCLVDNDPICVGENTVYRCVTRNGSAEDTNNVSLIL 480  
Db 421 SNGGTCSCAETTHWKGAAATHMCLVDNDPICVGENTVYRCVTRNGSAEDTNNVSLIL 480  
Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTUKGIAPGDARGEAILSS 540  
Db 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTUKGIAPGDARGEAILSS 540  
Qy 541 DTLTSPVSDTENTHYV 556  
Db 541 DTLTSPVSDTENTHYV 556

## RESULT 2

A86560  
60 kDa cysteine-rich omp [imported] - Chlamydomydia pneumoniae (strain J138)  
C:Species: Chlamydomydia pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
A:Accession: A86560  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: A86560  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-556 <STO>  
A:Cross-references: GB:BA000008; NID:98978928; PIDN:BAA98763.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: ompB  
C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 100.0%; Score 556; DB 2; Length 556;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKLIIRVVTVLALTSNACFASGGIEAAVAESLITKIVASAEKTPAPVPMPTAKKRVLR 60  
Db 1 MSKLIIRVVTVLALTSNACFASGGIEAAVAESLITKIVASAEKTPAPVPMPTAKKRVLR 60  
QY 61 RNKQPVOKSRGAFCDKEFYPCCEGRCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120  
Db 61 RNKQPVOKSRGAFCDKEFYPCCEGRCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120  
QY 121 YATVGSYPPIEILAIKCKDCVDVITQQLPCEAEFVSSDPETTPSGKLWIKIDRLGAG 180  
Db 121 YATVGSYPPIEILAIKCKDCVDVITQQLPCEAEFVSSDPETTPSGKLWIKIDRLGAG 180  
QY 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 240  
Db 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 240  
QY 241 EVWATGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCFQRRGQIT 300  
Db 241 EVWATGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCFQRRGQIT 300  
QY 301 NVATVTCYGGHKSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGLDLVLDVV 360  
Db 301 NVATVTCYGGHKSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGLDLVLDVV 360  
QY 361 IODTLPSSGVTVLEAPGEICCNVWRKEMCPGETLQFLVYKAQVPGRTNOAVTSE 420  
Db 361 IODTLPSSGVTVLEAPGEICCNVWRKEMCPGETLQFLVYKAQVPGRTNOAVTSE 420  
QY 421 SNGCTCTSCAETTHKNGLAATHMCVLDNDPFCVGENTVYRICVTRNGSAEDTNYSLIL 480  
Db 421 SNGCTCTSCAETTHKNGLAATHMCVLDNDPFCVGENTVYRICVTRNGSAEDTNYSLIL 480  
QY 481 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVFSVTLKGIAPGDARGEAILSS 540  
Db 481 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVFSVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHY 556  
Db 541 DTLTSPVSDTENTHY 556

## RESULT 3

B39439  
60K cysteine-rich outer membrane protein 1 precursor - Chlamydomydia psittaci  
C:Species: Chlamydomydia psittaci, Chlamydia psittaci  
C:Date: 21-Feb-1992 #sequence\_revision 27-Jun-1994 #text\_change 31-Mar-2000

C:Accession: B39439; SL2603  
R:Everett, K.D.E.; Hatch, T.P.  
J: Bacteriol. 173, 3821-3830, 1991  
A:Title: Sequence analysis and lipid modification of the cysteine-rich envelope prot.  
A:Reference number: A39439; MUID:91267949  
A:Accession: B39439  
A:Molecule type: DNA  
A:Residues: 1-557 <EVS>  
A:Cross-references: GB:M61116; NID:g144489; PIDN:AAB61619.1; PID:g144491  
A:Experimental source: strain 68C  
R:Watson, M.W.; Lambden, P.R.; Clarke, I.N.  
Nucleic Acids Res. 18, 5300, 1990  
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of  
A:Reference number: SL2603; MUID:90384851  
A:Accession: SL2603  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-44, 'A', 46-72, 'E', 74-557 <WAT>  
A:Cross-references: GB:X53512; NID:940625; PIDN:CAA37592.1; PID:g40627  
A:Experimental source: strain EAE/A22/M  
C:Function:  
O:Description: associated with the differentiation of reticulate bodies into elementary  
O:be an important virulence factor.  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-557/Product: 60K cysteine-rich outer membrane protein 1 #status predicted <MAT>  
Query Match 8.3%; Score 46; DB 1; Length 557;  
Best Local Similarity 100.0%; Pred. No. 27e-39;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 182 KCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 227  
Db 183 KCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 228  
RESULT 4  
JC5204  
60K cysteine-rich outer membrane protein 2 precursor - Chlamydomydia psittaci  
C:Species: Chlamydomydia psittaci, Chlamydia psittaci  
C:Date: 20-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 31-Mar-2000  
C:Accession: JC5204  
R:Hsia, R.; Bavoil, P.M.  
Gene 176, 155-162, 1996  
A:Title: Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: Stru  
A:Reference number: JC5203; MUID:97075924  
A:Accession: JC5204  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-558 <HSI>  
A:Cross-references: GB:U41759; NID:g1783376; PIDN:AAB41143.1; PID:g1783382  
A:Experimental source: strain GPIC  
C:Genetics:  
A:Gene: omp2  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-558/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>  
Query Match 8.1%; Score 45; DB 2; Length 558;  
Best Local Similarity 100.0%; Pred. No. 3e-38;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 183 CKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 227  
Db 185 CKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 229



R:Clarke, I.N.; Ward, M.E.; Lambden, P.R.  
Gene 71, 307-314, 1988  
A:Title: Molecular cloning and sequence analysis of a developmentally regulated cyste  
A:Reference number: JT0419; MUID:89138006  
A:Accession: JT0419  
A:Molecule type: DNA  
A:Residues: 30-547 <CIA>  
A:Cross-references: GB:M35148; NID:g144485; PIDN:AAA23119.1; PID:g144487  
A:Experimental source: serotype L1  
C:Genetics:  
A:Gene: omp2; omcB  
C:Function:  
A:Description: associated with differentiation of reticulate bodies into elementary b  
A>Note: essential for the structural integrity of the outer envelope of the elementar  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status experimental <PRO>  
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MA>

Query Match 4.5%; Score 25; DB 1; Length 547;  
Best Local Similarity 100.0%; Pred. No. 2.3e-17;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208  
Db 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 7  
B43584  
60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachoma  
C:Species: Chlamydia trachomatis  
C:Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 20-Aug-1999  
C:Accession: B43584; S13120; S18979; S24275  
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.  
F:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.  
A:Title: Immun. 59, 1196-1201, 1991  
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal  
A:Reference number: A43584; MUID:91147205  
A:Accession: B43584  
A:Molecule type: DNA  
A:Residues: 1-547 <DEL>  
A:Cross-references: EMBL:X54389; NID:g40763; PIDN:CAA38259.1; PID:g40764  
A:Experimental source: strain Bour, serotype E  
R:Coles, A.M.; Allan, I.; Pearce, J.H.  
Nucleic Acids Res. 18, 6713, 1990  
A:Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia  
A:Reference number: S13120; MUID:91067486  
A:Accession: S13120  
A:Molecule type: DNA  
A:Residues: 1-32, 'FT', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <COL>  
A:Cross-references: EMBL:X55503; NID:g40724; PIDN:CAA39396.1; PID:g40725  
A:Experimental source: strain DK20, serotype E  
C:Genetics:  
A:Gene: omp2; omcB  
C:Function:  
A:Description: associated with differentiation of reticulate bodies into elementary b  
A>Note: essential for the structural integrity of the outer envelope of the elementar  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MA>

Query Match 4.5%; Score 25; DB 2; Length 547;  
Best Local Similarity 100.0%; Pred. No. 2.3e-17;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208  
Db 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

R:Clarke, I.N.; Ward, M.E.; Lambden, P.R.  
Gene 71, 307-314, 1988  
A:Title: Molecular cloning and sequence analysis of a developmentally regulated cyste  
A:Reference number: JT0419; MUID:89138006  
A:Accession: JT0419  
A:Molecule type: DNA  
A:Residues: 30-547 <CIA>  
A:Cross-references: GB:M35148; NID:g144485; PIDN:AAA23119.1; PID:g144487  
A:Experimental source: serotype L1  
C:Genetics:  
A:Gene: omp2; omcB  
C:Function:  
A:Description: associated with differentiation of reticulate bodies into elementary b  
A>Note: essential for the structural integrity of the outer envelope of the elementar  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status experimental <PRO>  
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MA>

Query Match 7.0%; Score 39; DB 2; Length 554;  
Best Local Similarity 100.0%; Pred. No. 5.5e-32;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 489  
Db 449 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 487

RESULT 6  
A32244  
60K cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia tra  
C:Species: Chlamydia trachomatis  
C:Date: 12-Oct-1989 #sequence\_revision 27-Jun-1994 #text\_change 16-Jul-1999  
C:Accession: A32244; A43584; A36043; A30472; JT0419; S18981; S24277  
R:Allen, J.E.; Stephens, R.S.  
J. Bacteriol. 171, 285-291, 1989  
A:Title: Identification by sequence analysis of two-site posttranslational processing of  
A:Reference number: A32244; MUID:89123030  
A:Accession: A32244  
A:Molecule type: DNA  
A:Residues: 1-547 <ALL>  
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553  
A:Experimental source: strain L2/434/Bu  
A:Note: parts of this sequence, including the amino ends of the precursor and mature pro  
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.  
Infect. Immun. 59, 1196-1201, 1991  
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton  
A:Reference number: A43584; MUID:91147205  
A:Accession: A43584  
A:Molecule type: DNA  
A:Residues: 1-547 <DEL>  
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553  
A:Experimental source: serovar 2, strain L2/434/Bu  
R:Wahlberg, J.; Lundeborg, J.; Hultman, T.; Uhlen, M.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990  
A:Title: General colorimetric method for DNA diagnostics allowing direct solid-phase gen  
A:Reference number: A36043; MUID:90370827  
A:Accession: A36043  
A:Molecule type: DNA  
A:Residues: 294-402 <WAH>  
A:Experimental source: serotype L2  
R:Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.  
Gene 87, 105-112, 1990  
A:Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated transcr  
A:Reference number: JQ0514; MUID:90236284  
A:Accession: A30472  
A:Molecule type: DNA  
A:Residues: 1-46:528-547 <LAN>  
A:Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:g144485  
A:Experimental source: serotype L1

```

RESULT      8
D71515
60K cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia trachomatis
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence-revision 13-Sep-1998 #text_change 20-Aug-1999
C:Accession: D71515; C43584; S11673
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809
A:Accession: D71515
A:Molecule type: DNA
A:Residues: 1-553 <ARN>
A:Cross-references: GB:AE001317; GB:AE001273; NID:g3328863; PIDN:AC68042.1; PID:g332887
A:Experimental source: serotype D, strain UW-3/Cx
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M. Infect. Immun. 59, 1196-1201, 1991
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein
A:Reference number: A43584; MUID:91147205
A:Accession: C43584
A:Molecule type: DNA
A:Residues: 7-238 'V', 240-553 <DEL>
A:Cross-references: GB:X54388; NID:g40760; PIDN:CAA38257.1; PID:g40761
A:Experimental source: serotype C
R:Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N. FEMS Microbiol. Lett. 65, 293-297, 1989
A:Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence homology
A:Reference number: S11673
A:Accession: S11673
A:Molecule type: DNA
A:Residues: 7-553 <WAT>
A:Cross-references: EMBL:X53510; NID:g40681; PIDN:CAA37588.1; PID:g40683
A:Experimental source: serotype B
C:Genetics:
A:Gene: omp2; omcB
C:Function:
A:Description: associated with differentiation of reticulate bodies into elementary bodies
A>Note: essential for the structural integrity of the outer envelope of the elementary body
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-46/Domain: propeptide #status predicted <PRO>
F:47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>

```

```

Query Match      4.58; Score 25; DB 2; Length 553;
Best Local Similarity 100.08; Pred. NO. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 184 KITVWVKPLKGGCCFTAAATVCACPE 208
      |||
DB 181 KITVWVKPLKGGCCFTAAATVCACPE 205
      |||

```

```

Search completed: May 25, 2002, 22:26:30
Job time: 148 sec

```

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:26:32 ; Search time 12.65 seconds  
(without alignments)  
1701.824 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 556  
Sequence: 1 MSKLIRRVTVLALTMASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 12

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	100.0	556	1 OM6_CHLPN	P23700 chlamydia p
2	46	8.3	557	1 OM6_CHLPS	P23701 chlamydia p
3	25	4.5	547	1 OM6C_CHLTR	P26758 chlamydia t
4	25	4.5	547	1 OM6D_CHLTR	P18151 chlamydia t
5	25	4.5	547	1 OM6E_CHLTR	P23603 chlamydia t
6	25	4.5	547	1 OM6L_CHLTR	P21354 chlamydia t

ALIGNMENTS

RESULT 1

OM6\_CHLPN STANDARD: PRT; 556 AA.

AC P23700: Q9TQIG;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last annotation update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).

DE OM6C OR OMP2 OR CPN0557 OR CP0195.

GN Chlamydia pneumoniae (Chlamydophila pneumoniae).

OS Chlamydia pneumoniae (Chlamydiaceae: Chlamydiales).

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.

OX NCBI\_TaxID=83558;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IOL-207;

RC MEDLINE=90384850; PubMed=2402463;

RA Watson M.W., Al-Mahdawi S., Lamden P.R., Clarke I.N.;

RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207.";

RL Nucleic Acids Res. 18:5299-5299(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CWL029;

RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,

RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

RL Nat. Genet. 21:385-389(1999).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=AR39;

RC MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., L.,

RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

RA Eisen J., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MOpn and Chlamydia pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=J138;

RC MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

RA Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,

RA Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,

RA Ishii K., Shiba T., Hattori M., Kuhara S.;

RT "Comparison of outer membrane protein genes omp and pmp in the whole genome sequences of Chlamydia pneumoniae isolates from Japan and US.";

RL Nucleic Acids Res. 28:2311-2314(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=J138;

RA Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,

RA Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,

RA Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,

RA Ishii K., Shiba T., Hattori M., Kuhara S.;

RT "Comparison of outer membrane protein genes omp and pmp in the whole genome sequences of Chlamydia pneumoniae isolates from Japan and US.";

RL Nucleic Acids Res. 28:2311-2314(2000).

RN [6]

RP Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

CC -! FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES

CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL

CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT

CC VIRULENCE FACTOR.

CC -! SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X53511; CAA37590.1; -

CC EMBL; AE001640; AAD18697.1; -

CC EMBL; AE002180; AAF38068.1; -

CC EMBL; AP002547; BAA98763.1; -

CC EMBL; AB033786; BAA85939.1; -

CC PIR; S12602; S12602.

CC PHCI-ZDPAGE; P23700; -

CC TIGR; CP0195; -

CC InterPro; IPR003506; Chlam\_OMP6.

CC PRINTS; PR01336; CHLAMIDIOM6.

CC Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.

CC SIGNAL 1 22 POTENTIAL.

CC PROPEP 23 40

CC CHAIN 41 556 60 KDA OUTER MEMBRANE PROTEIN.

CC SEQUENCE 556 AA; 59719 MW; 8D7ED9234CC99458 CRC64;

Query Match 100.0%; Score 556; DB 1; Length 556;

Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;

Matches 556; Conservative 0; Mismatches 0;

QY 1 MSLKLRVTVLALTSMAFCASGIEAAVAESLITKIVASAETKPAVPMTAKKVLVR 60  
Db 1 MSLKLRVTVLALTSMAFCASGIEAAVAESLITKIVASAETKPAVPMTAKKVLVR 60  
QY 61 RNKQPVQKSGARCDKFEYPCBGRGQVPVBAQOQESYGRLYSVKVNDDCNVEICQSVPE 120  
Db 61 RNKQPVQKSGARCDKFEYPCBGRGQVPVBAQOQESYGRLYSVKVNDDCNVEICQSVPE 120  
QY 121 YATVGSPIEILAIAGKKDCVDVITQOLPCEAEFVSDPTTTSQGLKWKIDRLGAG 180  
Db 121 YATVGSPIEILAIAGKKDCVDVITQOLPCEAEFVSDPTTTSQGLKWKIDRLGAG 180  
QY 181 DKCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
QY 241 EYVNTGSAIARNYVDNPNVDPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQIT 300  
Db 241 EYVNTGSAIARNYVDNPNVDPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQIT 300  
QY 301 NVATVTCGGHKCSANVTTVNPPVPCQVNVISGADWSYVCKPVEYSISVSNPGLVLDHVV 360  
Db 301 NVATVTCGGHKCSANVTTVNPPVPCQVNVISGADWSYVCKPVEYSISVSNPGLVLDHVV 360  
QY 361 IQDTLPSGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
Db 361 IQDTLPSGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
QY 421 SNGCTCSAETTHWKGAAATMCMVLDNDPICVGNVTYRICVNTNRGSAEDTNVSLIL 480  
Db 421 SNGCTCSAETTHWKGAAATMCMVLDNDPICVGNVTYRICVNTNRGSAEDTNVSLIL 480  
QY 481 KFSKELQPIASSGPTKGTISGNVTVDALPKLGSRESVEFSYTLKGIAPGDARGAILSS 540  
Db 481 KFSKELQPIASSGPTKGTISGNVTVDALPKLGSRESVEFSYTLKGIAPGDARGAILSS 540  
QY 541 DTLTSPVSDTENTHY 556  
Db 541 DTLTSPVSDTENTHY 556

## RESULT 2

OM6\_CHLPS  
ID OM6\_CHLPS STANDARD; PRT; 557 AA.  
AC P23701;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).  
GN OMCB OR OMP2 OR ENVB.  
OS Chlamydia psittaci (Chlamydia psittaci).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83554;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EA/A22/M;  
RX MEDLINE=90384851; PubMed=2402464;  
RA Watson M.W., Lambden P.R., Clarke I.N.;  
RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia psittaci strain EAE/A22/M.";  
RL Nucleic Acids Res. 18:5300-5300(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=6BC;  
RX MEDLINE=91267949; PubMed=2050637;  
RA Everett K.D.E., Hatch T.P.;  
RT "Sequence analysis and lipid modification of the cysteine-rich envelope proteins of Chlamydia psittaci 6BC.";  
RL J. Bacteriol. 173:3821-3830(1991).  
RN [3]  
RP SEQUENCE FROM N.A.

RA Watson M.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBS). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC  
CC EMBL: X53512; CAA37592.1; -;  
DR EMBL: M61116; AAB61619.1; -;  
DR PIR: S12603; S12603.  
DR InterPro: IPR003506; ChlamOMP6.  
DR PRINTS: PR01336; CHLAMIDIAOM6.  
KW Outer membrane; Transmembrane; Signal; Virulence.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 40 POTENTIAL.  
FT CHAIN 41 557 60 KDA OUTER MEMBRANE PROTEIN.  
FT VARIANT 45 45 A -> S (IN STRAIN 6BC).  
FT VARIANT 73 73 E -> G (IN STRAIN 6BC).  
SQ SEQUENCE 557 AA; 59843 MW; 0D444F09EAA073C6 CRC64;

## Query Match

Best Local Similarity 8.3%; Score 46; DB 1; Length 557;  
Matches 46; Conservative 0; Mismatches 39;

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227  
Db 183 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 228

## RESULT 3

OM6C\_CHLTR  
ID OM6C\_CHLTR STANDARD; PRT; 547 AA.  
AC P26758;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 60 kDa outer membrane protein, serovar C precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).  
GN OMP2 OR OMP2B.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TW-3 / SEROVAR C;  
RX MEDLINE=91147205; PubMed=1997423;  
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;  
RT "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";  
RL Infect. Immun. 59:1196-1201(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Y.X., Caldwell H.D.;  
RX Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBS). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/embldb\\_license@isb-sib.ch](http://www.isb-sib.ch/announcement/embldb_license@isb-sib.ch)).

CC	of send an email to: <a href="mailto:cc@hawaii.gov">cc@hawaii.gov</a>
CC	EMBL; X54388; CAA38257.1; -
DR	EMBL; M85197; AAA23159.1; -
DR	PIR; C43584; C43584.
DR	InterPro: IPR003506; Chlam_OMP6.
DR	PRINTS; PR01336; CHLAMIDIAOM6.
DR	Outer membrane; Transmembrane; Signal; Virulence.
KW	POTENTIAL.
FT	SIGNAL 1 22
FT	PROPEP 23 40
FT	CHAIN 41 547
FT	SEQUENCE 547 AA; 58680 MW; 817BA5DC7FEA65D4 CRC64;
CC	60 KDA OUTER MEMBRANE PROTEIN, SEROVAR C.

Query Match 4.5%; Score 25; DB 1; Length 547;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

184 KITVWKPLKEGCCFATAATVCACPE 208  
175 KITVWKPLKEGCCFATAATVCACPE 199

RESULT	4				
OM6D_CHLTR		OM6D_CHLTR	STANDARD;	PRT;	547 AA.
ID		OM6D_CHLTR			
AC		P18151;			
DT		01-NOV-1990 (Rel. 16, Created)			
DT		01-NOV-1990 (Rel. 16, Last annotation update)			
DT		16-OCT-2001 (Rel. 40, Last annotation update)			
DE		60 kDa outer membrane protein precursor (Cysteine-rich outer membrane			
DE		protein) (CRP) (60 kDa cysteine-rich OMP).			
DE		OMCB OR OMP2 OR OMP443.			
GN		Chlamydia trachomatis.			
OS		Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
OX		NCBI_TaxID=813;			
RN		[1]			
RP		SEQUENCE FROM N.A.			
RC		STRAIN=B/JAL120/OT.			
RX		MEDLINE=90128208; PubMed=2612891;			
RA		Watson M.W., Lambden P.R., Ward M.E., Clarke I.N.;			
RT		"Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein:			
RL		sequence homology between trachoma and LGV biovars.";			
RL		FEMS Microbiol. Lett. 53:293-297(1989).			

RP SEQUENCE FROM N.A.  
RC STRAIN=B/TW-05/OT;  
RX MEDLINE=91141306; PubMed=2287277;  
RA Allen J.E., Carrone M.C., Beatty P.R., Stephens R.S.;  
RT "cysteine-rich outer membrane proteins of Chlamydia trachomatis  
RL display compensatory sequence changes between biovariants."; Mol. Microbiol. 4:1543-1550(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RN STRAIN=D/UW-3/CX;  
RC MEDLINE=99000809; PubMed=9784136;  
RX Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Ollinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RL Chlamydia trachomatis";  
RN Science 282:754-759(1998).  
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES  
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL  
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT  
CC VIRULENCE FACTOR.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.

39 -----

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55903; CAA39396.1; -.
DR EMBL; X54389; CAA38259.1; -.
DR EMBL; M85196; AAA23154.1; -.
DR PIR; S13120; S13120.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22
FT PROPEP 23 40
FT CHAIN 41 547
FT VARIANT 33 34 SL -> FT (IN STRAIN DK20).
FT VARIANT 121 121 I -> L (IN STRAIN DK20).
FT VARIANT 132 132 V -> A (IN STRAIN DK20).
FT VARIANT 458 458 N -> S (IN STRAIN DK20).
SQ SEQUENCE 547 AA; 58708 MW; 0520656084F4E20AB CRC64;

Query Match 4.5%; Score 25; DB 1; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.le-17; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199
|||||
|||||

RESULT 6
OMSL_CHLTR STANDARD; PRT; 547 AA.
AC P21354; P18586;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovars L1/L2/L3 precursor (Cysteine-
DE rich outer membrane protein) (60-kDa CRP).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=L2/434/80; PubMed=2914847;
RX MEDLINE=89123030; PubMed=2914847;
RA Allen J.E., Stephens R.S.;
RT "Identification by sequence analysis of two-site posttranslational
RT processing of the cysteine-rich outer membrane protein 2 of Chlamydia
RT trachomatis serovar L2."
RL J. Bacteriol. 171:285-291(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=404 / SEROVAR L3;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
RA Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative
RT 15-kilodalton protein between the trachoma and lymphogranuloma
RT venereum biovars of Chlamydia trachomatis."
RL Infect. Immun. 59:1196-1201(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=L1/440/LN;
RX MEDLINE=89138006; PubMed=3066701;
RA Clarke I.N., Ward M.E., Lambden P.R.;
RT "Molecular cloning and sequence analysis of a developmentally
RT regulated cysteine-rich outer membrane protein from Chlamydia
RT trachomatis."

```

```

RL Gene 71:307-314(1988).
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23001; AAA23152.1; -.
DR EMBL; X54390; CAA38261.1; -.
DR EMBL; M35148; AAA23119.1; AUT_INIT.
DR PIR; A32244; A32244.
DR PIR; S18981; S18981.
DR PIR; A43584; A43584.
DR PIR; JTO419; JTO419.
DR Siena-2DPAGE; P21354; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22
FT PROPEP 23 40
FT CHAIN 41 547
FT SEQUENCE 547 AA; 58782 MW; 78CEB41CCF98472D CRC64;

Query Match 4.5%; Score 25; DB 1; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.le-17; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199
|||||
|||||

Search completed: May 25, 2002, 22:30:06
Job time: 214 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2002, 16:42:13 ; Search time 4990.55 Seconds  
(without alignments)  
8851.914 Million cell updates/sec

Title: US-09-523-647-1  
Perfect score: 2111  
Sequence: 1 ttgtacaggtagtaggaga.....catcaaatctacatcccgta 2111

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: gb\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
-----						

1	2111	100.0	2711	1	CPCRMP
2	2107.8	99.8	11938	1	AE002180
c	2107.8	99.8	300550	1	AP002547
c	2106.2	99.8	10422	1	AE001640
5	1671	79.2	1671	6	AX349525
6	1185.2	56.1	3650	1	CHTCRPNV
7	1185	56.1	1185	1	AF111201
8	1183.6	56.1	2478	1	CPSCROMP
9	1158.6	54.9	8950	1	CPA1759
10	1156.6	54.8	3186	1	AF240773
11	1031.8	48.9	1604	1	CPU76760
12	944.8	44.8	1604	1	CPU76761
13	889.6	42.1	2390	1	CHTOMPE
14	888	42.1	2385	1	CTCROMP
15	888	42.1	2390	1	CHTOMPA
c	887.6	42.0	11141	1	AE001317
17	887.6	42.0	1734	1	AF304326
18	887.6	42.0	1738	1	AF304327
19	887.6	42.0	1752	1	AF304331
20	887.2	42.0	1753	1	AF304328
21	884	41.9	1753	1	AF304330
22	882.8	41.8	1919	1	CHTOMPA2A
23	881.4	41.8	1753	1	AF304329
24	877.6	41.6	2345	1	CTSCCRP
25	876.8	41.5	3012	1	CHTCRPA
c	876.6	41.5	12442	1	AE002341
27	876.4	41.5	1726	1	CTOMP2
28	876	41.5	2342	1	CTSECRP
29	864.8	41.0	2430	1	CTSL3CRP
30	817	38.7	1612	1	AF304332
31	795.8	37.7	1185	1	AF111200
c	767.6	36.4	1947	6	AX338361
33	725.8	34.4	1185	1	AF111199
c	719.6	34.1	1278	6	AX338362
34	549.4	26.0	551	1	CPU56926
35	544.6	25.8	551	1	CPU56925
36	523.2	24.8	528	1	AF347609
37	523.2	24.8	386	1	AF102831
38	379.6	18.0	386	1	AX338251
39	363.6	17.2	731	6	AF367407
40	287.8	13.6	548	1	CPU56924
41	284	13.5	534	1	CPU56927
42	282.4	13.4	534	1	AF367405
43	231	10.9	471	1	AF367406
44	231	10.9	471	1	AF367406
45	228.4	10.8	230	6	AR144060

#### ALIGNMENTS

RESULT	1	CPCRMP	2711 bp	DNA	linear	BCT 06-NOV-1996
LOCUS	Chlamydia pneumoniae gene for cysteine rich outer membrane protein.					
DEFINITION	X53511					
ACCESSION	X53511.1	GI:550564				
VERSION						
KEYWORDS	outer membrane protein.					
SOURCE	Chlamydia pneumoniae.					
ORGANISM	Chlamydia pneumoniae					
REFERENCE	1 (bases 1 to 2711)					
AUTHORS	Watson, M.W.					
TITLE	Direct Submission					
JOURNAL	Submitted (22-JUN-1990) Watson M.W., Southampton University England, Department of Microbiology, Faculty of Medicine, South Lab and Path block, Southampton General Hospital, Tremona road, Southampton, SO9 4XJ, UK					
REMARK	revised by [4]					
REFERENCE	2 (bases 1 to 1800)					
AUTHORS	Watson, M.W., al-Mahdawi, S., Lamden, P.R. and Clarke, I.N.					
TITLE	The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207					
JOURNAL	Nucleic Acids Res. 18 (17), 5299 (1990)					

MEDLINE 90384850  
REFERENCE 3 (bases 1 to 2711)  
AUTHORS Watson,M.W.  
TITLE Direct Submission  
JOURNAL Submitted (09-JAN-1991) M.W.Watson, Univ., Department of  
Microbiology, Southampton General Hospital, Southampton, S09 4XY,  
UK  
REMARK revised by [4]  
REFERENCE 4 (bases 1 to 2711)  
AUTHORS Watson,M.W.  
TITLE Direct Submission  
JOURNAL Submitted (03-FEB-1993) M.W.Watson, Univ., Department of  
Microbiology, Southampton General Hospital, Southampton, S09 4XY,  
UK  
REFERENCE 5 (bases 1 to 2711)  
AUTHORS Watson,M.W., Clarke,I.N., Everson,J.S. and Lambden,P.R.  
TITLE The Crp operon of Chlamydia psittaci and Chlamydia pneumoniae  
JOURNAL Microbiology 141 (Pt 10), 2489-2497 (1995)  
MEDLINE 96036206  
FEATURES Location/Qualifiers  
1..2711  
/organism="Chlamydia pneumoniae"  
/strain="IOL-207"  
/db\_xref="taxon:83558"  
prim\_transcript 204..2571  
/gene="60kDa Crp"  
prim\_transcript 204..643  
/gene="9kDa Crp"  
gene 204..645  
/gene="9kDa Crp"  
gene 204..2644  
/gene="60kDa Crp"  
prim\_transcript 204..2509  
/gene="60kDa Crp"  
prim\_transcript 204..2507  
/gene="60kDa Crp"  
prim\_transcript 204..2508  
/gene="60kDa Crp"  
prim\_transcript 204..644  
/gene="9kDa Crp"  
prim\_transcript 204..645  
/gene="9kDa Crp"  
prim\_transcript 204..2572  
/gene="60kDa Crp"  
prim\_transcript 204..2573  
/gene="60kDa Crp"  
prim\_transcript 204..2574  
/gene="60kDa Crp"  
prim\_transcript 204..2575  
/gene="60kDa Crp"  
prim\_transcript 204..2641  
/gene="60kDa Crp"  
prim\_transcript 204..2642  
/gene="60kDa Crp"  
prim\_transcript 204..2643  
/gene="60kDa Crp"  
prim\_transcript 204..2644  
/gene="60kDa Crp"  
prim\_transcript 204..2570  
/gene="60kDa Crp"  
RBS 291..297  
/gene="9kDa Crp"  
CDS 300..572  
/gene="9kDa Crp"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA37589.1"  
/db\_xref="GI:550565"  
/db\_xref="SPTREMBL:O46183"  
translation="KKAVLIAMFCGVVSLSSCRIIVDCCFEDPCAPSSCNPCPEVIR  
KKERSCGNAGSYVPSNFCGTECSNQSQPVKGCYSPDGRCKQ"  
728..732  
RBS  
/gene="60kDa Crp"

CDS  
739..2409  
/gene="60kDa Crp"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA37590.1"  
/db\_xref="GI:550566"  
/db\_xref="SWISS-PROT:P23700"  
translation="MSKLIRRVTVLALTSMA SCFASGSGIEAAVAESLTKIVASAE  
T  
KPAVPMTAKRVLRRNKQPVQKSRGAFCDKEFYPCIEGRQCPVQAOQESCYGRLY  
SVKVNDCNVEICQSPYVATVGPPIEILAIKKDCVDVVITQOLPCEAEFVSSDP  
ETTPTSQKLVNKIDRLGAGDKCKITVHVKPLKECCCTAATVACAPELRSTKCGQP  
AICLQEGPDCACLRCPYCIKVNTSALARNVTVDNPNVDPGVSASGQRLVSNL  
GDMRPGKRVETVFCFNPGRGDIQDITPLSPGVTVLEAPGECICCNVKNRIKEMC  
WSYVCKPVEYSISVSNPDGLVLDHVDITLPGVTVLEAPGECICCNVKNRIKEMC  
PGETLQFKLVVKAQVPGRTNQAVTSENGCTCTCAETTHMKGLAATHMCLVDTN  
DPCICGNTVYRICVTRNGSAEDTNVSLILKFSKELQPIASSGPKGTISGNTVVVFD  
LPKLGSKESVEFSVTLKGIAPCDARGEAILSSDILTSPVSDTENTHVV"

BASE COUNT 824 a 502 c 546 g 839 t  
ORIGIN  
Query Match 100.0%; Score 2111; DB 1; Length 2711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ttgatacaggtagtaggagatgaataatctcctgactacctaattcagataataaaccc 60  
Db 601 TTGATCAGGTAGTTAGGAGATGAATTAATTCCTGACTACCTTTAATTCAGATAATAAACCC 660  
Qy 61 aaatgttgagggtagagttttacaaaacattctacccatgagcagagaagaaaaataaac 120  
Db 661 AAATGTTGAGGGTAGAGCTTTACAAAACATTCATCCCATGGCAGAGAAAAATAAAC 720  
Qy 121 atgcgataaggagatccctatgtccaaactcatcagacagtagttacggtcctctgcgcta 180  
Db 721 ATGGCATAGGAGATCCCTATGTCCAAACTCATCAGCAGTAGTTACGGTCTCTCGCGTA 780  
Qy 181 acgagtagcgagttgcttgcagcggggtagagggcgctgtagcagagatctctg 240  
Db 781 ACGAGTAGCGGAGTTGCTTTGCCAGCGGGGTATAGAGGCGCTGTAGCAGAGTCTCTG 840  
Qy 241 attactaagatcgtcgctagtcgcgaaacaaagcagcacctgttccatgacagcgaag 300  
Db 841 ATTACTAAGATCGTCGCTAGTGGGAAACAAAGCCAGCACCTGTCTCTATACACCGAAG 900  
Qy 301 aaggttagactgtccgttagaataaaacacagttgaacaaaaaacgctgggtctttt 360  
Db 901 AAGTTAGACTTGTCCGTAGAAATAAACACACAGTTGAACAAAAAGCCGTGCTTTT 960  
Qy 361 tgtgataaagaattttatccctgtgagagggagcagtagtcaacctgtagaggtcagcaa 420  
Db 961 TGTGATAAAGAAATTTATCCCTGTCAAGAGGAGCATGTCAACCTGTAGAGGCTCAGCAA 1020  
Qy 421 gactctgtcacggaagattgtattctgtaaaagtaaacgatgatgtgcaacgtgagaatt 480  
Db 1021 GAGTCTTGCTACGGAGATGTATTCTGTAAAGTAAACGATGATTGCAAGTAGAAATT 1080  
Qy 481 tgcagtcctgtccagaatacgtactgttaggtatcctctaccctattgaaatcctgtct 540  
Db 1081 TGCAGTCCCTTCCAGAAATACGCTACTGTAGGATCTCTTACCCCTATTGAAATCCTTGT 1140  
Qy 541 ataggcaaaaaagattgt 600  
Db 1141 ATAGGCAAAAAAGATTGT 1200  
Qy 601 ttgcgtaagcagtagatccagaaacaaactcctacaagtgatgggaaatagtcgtgaaatc 660  
Db 1201 TTGTAAGCAGTAGTATCCAGAAACAACTCTTACAGGTGATGGGAAATTAGTCTCGAAAAATC 1260  
Qy 661 gatcgctgggtgcaggagataaatgcaaaattactgtatgggtgaaacacctctaaagaa 720  
Db 1261 GATCGCTGGGTCCAGGAGATAAATGCAAAATTACTGTGTGGGTAAACCTCTTAAGAA 1320



721 ggttgctcttcacagctgctactgtatgtgtcttgccagagctccctgtttatactaaa 780  
1321 ggttgctgcttcacagctgctactgtatgtgtcttgccagagctccctgtttatactaaa 1380  
781 tgcggtcaaccagcatttatttaagcaagaagacactgactgtctgctgaagatgc 840  
1381 tgcggtcaaccagcatttatttaagcaagaagacactgactgtctgctgaagatgc 1440  
841 cctgtatgctacaaatcgaaagtagtagaacacagagatctgctattgcccgttaacgaact 900  
1441 cctgtatgctacaaatcgaaagtagtagaacacagagatctgctattgcccgttaacgaact 1500  
901 gtatgataatcctgttcccgatggctattctcatgcatctggtcaagagattctctcttt 960  
1501 gtatgataatcctgttcccgatggctattctcatgcatctggtcaagagattctctcttt 1560  
961 aacttaggagacatagacactgctgataaaaaggattttacacagttgagttctgcctcaaa 1020  
1561 aacttaggagacatagacactgctgataaaaaggattttacacagttgagttctgcctcaaa 1620  
1021 agagaggtcaaatcactaactgctgctgctgctgctgctgctgctgctgctgctgctgct 1080  
1621 agagaggtcaaatcactaactgctgctgctgctgctgctgctgctgctgctgctgctgct 1680  
1081 gcaaatgtaactacagttgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1140  
1681 gcaaatgtaactacagttgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1740  
1141 tggcttactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1200  
1741 tggcttactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1800  
1201 gttcttactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1260  
1801 gttcttactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1860  
1261 cctggtgagagatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1320  
1861 cctggtgagagatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1920  
1321 gaaacccctcagtttaaaactgtagtgaagctcgaagttcctctggaagattcacaaataca 1380  
1921 gaaacccctcagtttaaaactgtagtgaagctcgaagttcctctggaagattcacaaataca 1980  
1381 gttgcagtaactgtagtgaagctcgaagattcgaagattcgaagattcgaagattcgaagatt 1440  
1981 gttgcagtaactgtagtgaagctcgaagattcgaagattcgaagattcgaagattcgaagatt 2040  
1441 cattgaaaggtcttcagctaccatctgctgctgctgctgctgctgctgctgctgctgctg 1500  
2041 cattgaaaggtcttcagctaccatctgctgctgctgctgctgctgctgctgctgctgctg 2100  
1501 gtagggaataactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1560  
2101 gtagggaataactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2160  
1561 aacgtatcttaactctgaagttctcaaaagaacttcagccaatagcttcttcaggtcca 1620  
2161 aacgtatcttaactctgaagttctcaaaagaacttcagccaatagcttcttcaggtcca 2220  
1621 actaaaggaacgatttcaggttaacccgttcttcagacgttttaactaaactcgttct 1680  
2221 actaaaggaacgatttcaggttaacccgttcttcagacgttttaactaaactcgttct 2280  
1681 aagaaatctgtagagttttctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1740  
2281 aagaaatctgtagagttttctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2340  
1741 gaagctattcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1800  
2341 gaagctattcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2400  
1801 gttgataaattcttaaggaattatcttaagcagagcgatattccgctctgcttaggat 1860

2401 GTCTATTAAATCTAAGGAATTATCTAAGCAGACCGATATTCGCTCTGCTTTAGGAT 2460  
1861 agcttcaagaagtagtaccgcttttagtaccttaactaaagcggttttttggctttata 1920  
2461 AGCTTTCAAAGAGTAGTACCGCTTTAGTACCTTACGTACTAAAGCGGTTTTTTTGTATA 2520  
1921 agctcttcaatccaatcgtagagtttcttaatacaagataattttaagtttctgaaatc 1980  
2521 AGCTCTTCAATCCAATCGTAGAGTTCTTAAATCAAGATATTTAATTTAAGTTTCTGAAATC 2580  
1981 ctaagatt 2040  
2581 CTAAGATTTATTTAAAGCCCATCTTTTAGGTATCTAATTAATAATTTTAAATTAAGCT 2640  
2041 ttctctagtgtaacctgcttcttttaggaactacactaggaacggtatgcatcaaatc 2100  
2641 TTTCTAGTGTAACTGCTTCTTTTAGGAACCTACACTAGGAGACGGTATGTCATCAATC 2700  
2101 tacatccgcta 2111  
2701 TACATCCCGTA 2711  
RESULT 2  
AE002180 11938 bp DNA linear BCT 30-MAY-2000  
LOCUS Chlamydomophila pneumoniae AR39, section 16 of 94 of the complete genome  
DEFINITION AE002180 AE002161  
ACCESSION AE002180.2 GI:8163382  
VERSION  
KEYWORDS  
SOURCE Chlamydomophila pneumoniae AR39.  
ORGANISM Bacteria: Chlamydiales; Chlamydiaceae; Chlamydomophila.  
REFERENCE 1 (bases 1 to 11938)  
AUTHORS Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.  
TITLE Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)  
MEDLINE 20150255  
PUBMED 10684935  
REFERENCE 2 (bases 1 to 11938)  
AUTHORS Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
COMMENT On Jun 1, 2000 this sequence version replaced gi:7189117.  
FEATURES Location/Qualifiers  
source 1..11938  
/organism="Chlamydomophila pneumoniae AR39"  
/strain="AR39"  
/db\_xref="taxon:115711"  
/note="synonym: Chlamydia pneumoniae AR39"  
73..1839  
/gene="CP0187"  
73..1839  
/gene="CP0187"  
/note="similar to GB:M54884 SP:P21893 PID:147550  
PID:887842 GB:U00096; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="single-stranded-DNA-specific exonuclease RecJ"  
/protein\_id="AAF38060.1"

```

/db_xref="GI:7189118"
/translation="MTNSDNAAGLLWAHPKEDPAFLGMIIEKHELPPTVAQIFISR
GFQTIOEHKFLYSHLSLYDPGLFDMKKAVERLLLRKEHVMYIGDSVDGDMTG
VALLVEFRDIDVHVSYFELGAILKQHSETSLIAKKEEGITLLITVDCGTAGKEV
SDITRQGDVILVTDHMPGTPIPHCVATLNPRLDRTYNNRELTGVGVAFLARGVLN
ALISNVPKSGSLKULDLVTGTTIDVGVLLGENVMVRYGKETARGARGLNK
ICALCGVEKSESTDIVLKAPKNSLGRDDPAKGVLELLTODDERVDALIMELDN
INRQRLEAEVQDQVEILNSPEILKQAAIVLSSTAWHARVPIISARLAKYNKP
VVIATIQGIGKSGARTIGSPFLVLKCSLLLSYGGHFAAGVIMKEDKVDFFKK
KFVHLVNSSLKGGDTLPHLEIDAVDEDAIDYDILLASMELEFPFGKGLMPIFYSKVR
QVRYPKVLPGNHLKLYLSOKERNLEGVAFGLGRHADALKASWHYPLEIATYPRLSQTS
GSGVHLLVRDRISSEPRSD"
2093..3214
/gene="CP0188"
CDS
2093..3214
/gene="CP0188"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF38061.1"
/db_xref="GI:7189119"
/translation="MSAIAREQHOAAILDMHPKPSIAMSSESQARTSWEKROAHPYLY
RLEIITGVVFKLGLIFFTPLGLFWLOKICNFILGAGGWIFRPICRDSNLLRQA
YAALFSGFDHVSRRVRCLOYDEVDFDGLRLUPNAKPKRMILISNGNSDCLERY
TVLQGEKDFIRTAESQSNILFNTPGVKSGNITRNNVKSYQACVRYLRDEPAG
POARQIVAGYSLGASQAELSKETADGSDSVRFVVKDGRGARSTCAVAKQFISGLG
VMLANTHNNINSEKSKOLHCPFLFYIGKDSQGNLIGDLGFKKETCFAPFLDPKNL
ECSGKKIPVAQTGLRHHDHILSDDDVIKEVAGHRRFDN"
3725..4261
/gene="CP0189"
CDS
3725..4261
/gene="CP0189"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF38062.1"
/db_xref="GI:7189120"
/translation="MACOHEGCVLELERIEDIKSDTKWVSITQAAKLHNVTROA
IYVAKOKKLKASKETREIDIKLEEKRRYRKRKSLYQGLFVDFNGKGCYSINOV
AQIIGIPQVKYVYATRTGIRGERKGAAWIHVSEIERKYNEYLKQAAKKLKGAEPK
EHOAPNEPPEIPEFESN"
4517..6034
/gene="CP0190"
CDS
4517..6034
/gene="CP0190"
/translation="similar to PID:1783380; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA synthetase"
/protein_id="AAF38063.1"
/db_xref="GI:7189121"
/translation="MWNENVRVAPSPGDPHYGTAYMALFNEIFAKRFGKGMILRI
EDTDRTRQDVEENFESALRWCGIOWRDPGVGGPIGPIROSERTKIYQGYVETLLK
TDCAYKCEATPOELAEWRVASTLGRGTDRIYRYLSPEEVSAREAAQGPYITRLKV
PLSGECVFEDYSGKRVYFPWADVDQVLVKSDGFPYHFANVIDDLMLGITHVLRGEF
NLSSTPKHLLLEAFGEWPEPFLHMLLNPDTGKLSRKNPTSIFYRDSGVKKEAF
VNFLLTGYSMGDEEVSLERIETFPNRRIGKSGAVFDIOKLDMNKHLYLNHESGP
ECLLLEQGLMNDDEFFLKLPLCQSRITTLAEFINLTSFSGLLERYVBELLPOAL
SPEKAAILLYSVYKYLEKTQDWTKECYLGSKLQAQFNVHHKAILPLLYVAITGCK
QGLPLFDSIEILGKPRARALVYAEKLLGVGPKLAATVDRKMFQREDFEATFDL"
complement(6199..6387)
/gene="CP0191"
CDS
complement(6199..6387)
/gene="CP0191"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11

```

```

/product="conserved hypothetical protein"
/protein_id="AAF38064.1"
/db_xref="GI:7189122"
/translation="MKLFYFCNIIFLTSLFSVSCHTTTPPRSYILAQGRTPLVNKAHS
LEFFLARVSFNTCYNTNL"
6645..6752
/gene="CP0192"
CDS
6645..6752
/gene="CP0192"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF38065.1"
/db_xref="GI:7189123"
/translation="MFSKKSFNKKIYLEKIIDKIFIKIIFICNIVHGNVLY"
6829..7101
/gene="CP0193"
CDS
6829..7101
/gene="CP0193"
/translation="similar to PID:1783381; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="cysteine rich outer membrane protein 3"
/protein_id="AAF38066.1"
/db_xref="GI:7189124"
/translation="MKKAVLIAAMFCGVVLSGCCRIVDCFCFEDPCAPSSCNPCEVIR
KKERSGGNACGYSVPSCNPGCSQSPQVKGCTSPDGRCKQ"
complement(7165..7281)
/gene="CP0194"
CDS
complement(7165..7281)
/gene="CP0194"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF38067.1"
/db_xref="GI:7189125"
/translation="MSLDIGISYRMFIFFCHRVCFVNSYPOLHLLSELNR"
7268..8938
/gene="CP0195"
CDS
7268..8938
/gene="CP0195"
/translation="similar to SP:P23700 PID:550566 PID:1326243
PID:1326245 GB:AE001363; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="60 kDa outer membrane protein"
/protein_id="AAF38068.1"
/db_xref="GI:7189126"
/translation="MSKLIIRVVTVLALTSMA3CFASGGIEAAVAESLITKIVASAET
KPAVPMATKRLVRNRKQPFQKSRGAFCEFYPCEEGRCPVPAQOQESCVGRLY
SVNDDCNVEICQSVPEYATVGSPIEILLAGKDCVDVWITQOLPCAEVSSDP
ETPTTSDGLKWKIDRLGAYGKCIITWVKPLKGGCCFTAAACAPELKSYTKCGOP
AICIKQEGDPDCAICLRCPVCYKIEVNTGSAIARNVDPNVPDGTSHASQORVLSFNL
GMRPDKKVFYVEFCORRGQITNTVATVTCGGHKCSANVTTVVNEPCQVQVNSGAD
WSYVCKPVEYSISVNPGLVLDHVVQDTLVSGTCTSCAETTHWKGLAATHMCLVDTN
PCETLOFKLVKRAQVGRFTNQVAVTSESNGTCTSCAETTHWKGLAATHMCLVDTN
DPICGENTVYRIVTNRGSAEDTNVSLILKFSKELOPIASSGPTTKGTISGNTVVFDA
LPLKGSKEVSFEFSVTLKGIAPGDARGEAILSSDTLTSVPSDENTHNY"
9217..9807
/gene="CP0196"
CDS
9217..9807
/gene="CP0196"
/translation="similar to SP:P18587; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="15 kDa cysteine-rich outer membrane protein"

```



```

Db 9110 CTAAGATTATTTTAAAGCCCATCTTTTATAGGTATGTAATTAATAATTTTAAATGCT 9169
Qy 2041 ttctcctagtgtaacctgtctcttttaggaacacactaggaacggtatgtcatcaaatc 2100
|||||tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Db 9170 TTTCCCTAGTGTAACCTGCTTCTTTAGGAACATACACTAGGAGAACGGTATGTCATCAAAATC 9229
Qy 2101 tacatcccgta 2111
|||||tttttttttt
Db 9230 TACATCCCGTA 9240

RESULT 3
AP002547/c
LOCUS
DEFINITION
Chlamydomophila pneumoniae J138 genomic DNA, complete sequence,
section 3/4.
ACCESSION
AP002547 AB0333786 AB0333787 AB0333816 AB033817 BA000008
VERSION
AP002547.2 GI:10176693
KEYWORDS
Chlamydomophila pneumoniae J138 (strain:J138) DNA.
SOURCE
Chlamydomophila pneumoniae J138
Bacteria: Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE
1 (sites)
Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Kishi,F., Ouchi,K.,
Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.
Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
20330349
2 (bases 1 to 300550)
Shirai,M.
Direct Submission
Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
School of Medicine, Department of Microbiology; 1-1-1
Minamikuoshi, Ube, Yamaguchi 755-8505, Japan
(E-mail:shirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,
Fax:81-836-22-2415)
On Sep 15, 2000 this sequence version replaced gi:6172298
91:6172300 gi:6172396 gi:6172398 gi:8978869.
AB033786-AB033787, AB033816-AB033817: Submitted (25-Oct-1999).
FEATURES
source
1..300550
/organism="Chlamydomophila pneumoniae J138"
/strain="J138"
/db_xref="taxon:138677"
/note="synonym:Chlamydia pneumoniae (strain J138)"
complement(21..794)
/gene="dapP"
complement(21..794)
/gene="dapP"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translation="MAFVSESTISFYIYSGAGNRFLLGETLPEYEDVRFLEQETRVLD
GFLYLKPSCADAGLIIFNSDGRPTMCGNLCALAHLSQKGSIDSVSVDLSGLYS
GYFVSWDRVLVMDLADWRASVHRLESRPDLPKVCVCIHTGVHVVILPEISTDL
SLIGFLRYHQTSPDGVNVNVQILHCGQLRVYERGVCEGTAACTGALASALVV
SNSYCKRESIQHTWGGELMTVSNRGRVYLGQSVTRDL"
complement(763..1338)
/gene="clpP.1"
complement(763..1338)
/gene="clpP.1"
/codon_start=1
/transl_table=11
/product="Clp protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translation="MADGEVHKLRDIIEKELLRARVFESEPTSEKASDAIKKLWYL
ELKDPGKIVFVINSFGSVVDAGFVMDQIKMLTSPVTVWVTLAASMGSVLSLCAAP
GRFATPSRIMIHOPISGIPITGQATDLDIHAREILTKARIIDVVEATNQPDII
EKAIDRDMMWMTANEAKDQGLLDGLILSFNDL"

```

---

```

gene
CDS
complement(1358..2917)
/gene="glyA"
complement(1358..2917)
/gene="glyA"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translation="MLKVEPKPKFAIVEIFTKVVAVVUSLLLUKLENASGKKGOSLAS
TAYLAALDHLNAPFSGERLIDELKSORSHLMIASENYSSLSVLAMGNLLTDKYC
EGSPKRFKYSCCENVDATMECVETAKELFAADCAVOPHSGADLANLVAULTHKV
KCPYDYNPOTGCFDFAEISRLAKYKPKVLIAGYSSYRRLNFAVLKQIATREYESTLNA
WDMAHFAGLVAGVGVDEENPIYADIIVTTTHKTLRGPRGLVLAATREYESTLNA
CPLMMGGPLPHVIAAKTALKALSYDFKFAHOVNVNNAERLAEFLSHGLRLTGST
DNHMMVLDLGSLSGKIAEDILSSVGIANNRSLPSDAIGKWDTSGLRGLTALTTL
CMGIDEMEVAADIIVKVLNRILRILSCHVEGSKKNGKELPEATAOEARDRVRNLLLRFP
LYPEIDLEALV"
3107..3778
/gene="CP10522"
3107..3778
/gene="CP10522"
/codon_start=1
/transl_table=11
/product="CT433 hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="MTLYLGLNQTKARKYOAHYLPILTLPYAKSTPONKRALQFLPO
ATHVLTSPSSTHLFSRMTSLLSKATLTKTYLCIGESTKERLLSLFGQVKKVYVATQ
BIAGKIFPLQAQPSARILYPHSSILARPVIREFLYNRFTFFSPHYTVKPRKLNKI
LSYKKIKIIFTPSTVRAFAKIFPRFPKTYMQGRMTLQEFQKFSQKQVSLLETLGK
SRTSP"
3843..4175
/gene="CP10523"
3843..4175
/gene="CP10523"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATPGDGTAPSLPPRPRYNFKLALFVTLAIALWIAL
IATTIATGCTHPLCSPILTALPLFYISRYICSHYARNVYIALDVDPDHKLODMKS
HSPIFSQR"
4223..5302
/gene="CP10524"
4223..5302
/gene="CP10524"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSPGSRTESSQSVLSVYPRKEIAPKQFTIAKLTALILASL
ALGALVAGISITIVLGNPVELALLITLTFVVTFLVYHOMTSKYSSNMQVLEONPK
PLGKAMGKNVDCYSNEMQYNNHLPKFKVAIQDASQEQPTFLGLRVIEKNQST
GIIFNPVPTNIDNTATNLSTILYSLTKDKSVMDTKQKQGGKAGDEDPSEPTFVRV
VKLEALDQFLNLSAEEKSKILPTFLGHVCGPKSELNQOQYVQALLAYENCL
KAATESHAAIVALPLFTSVVEVPPEILPKEGTFYWDNTOAFCKRALLDAIQNTALR
YQSRSLVILQDPENTIESQRSE"
5642..6406
/gene="CP10525"
5642..6406
/gene="CP10525"
/codon_start=1
/transl_table=11
/product="CT398 hypothetical protein"
/protein_id="BAA98731.1"
/db_xref="GI:8978896"
/translation="MHDALSLIAIOELDTKIMLRMRVKKEHQELAKVQSLKSDIRR
KVOEKELEMENLKTQIRDGNRIQETSEQINKLENOAAVKKMDEFNALTQEMTTANK
ERRSLHQLSLMDKQAGGDLIVLSKESLASTENSSSVIEKEIFESIKKINEEGKAL

```

LEQTELKHA<sup>2</sup>NPPELLSIYTERLLNNKKDKRVVYP<sup>2</sup>ENRVCSGCH<sup>2</sup>VLTPQHENLVRRKK<sup>2</sup>  
 RLIFCEHCSRLIYQESQVNAQNSAKRRRRRAAV<sup>2</sup>  
 /gene="kpsf"  
 CDS complement(6827. .7816)  
 /gene="kpsf"  
 /codon\_start=1  
 /transl\_table=11  
 /product="GutQ/KpsF family sugar-P isomerase"  
 /protein\_id="BAA98732.1"  
 /db\_xref="GI:8978897"  
 /translation="MPSPMISTDVQDILHGKQKAVDPFFQAFQPKAMOLAELKILGH  
 SGWFFSGVSGKSCVARKLIVATLQISLRAELFFSPVDLLHGDLGLVSGDILVCLFSKS  
 TQOMIEDLVPHLKRRRILVATLSMPYSLNLAASDLVVLIPSPVMEKDPENLLPTSTG  
 GTCMIFDGLMILLFRRALISVLTSGKNHPSQVGMKANGKVPKPEKTEVPLTSTG  
 DKYFSILVESAYCCGCVLQVDPQRLMCIITDDLRSLASGYGEVLSLSLEKVMTA  
 NPACITEDSDIATLQIMRESSPVAVLVDLNEENRHVIGLHMHITLAKAGLL"  
 complement(7843. .9024)  
 /gene="sucB\_2  
 CDS complement(7843. .9024)  
 /gene="sucB\_2"  
 /codon\_start=1  
 /transl\_table=11  
 /product="dihydrolipoamide succinyltransferase"  
 /protein\_id="BAA98733.1"  
 /db\_xref="GI:8978898"  
 /translation="MIFEFRRPKIGETSSGGSIVHWLKNLGDHVARDEPLIEVSTDKI  
 ATELPSPKAGRLVRFVCNEGDEVASGDVGLGLEIFEISEADDESTCPLTSEKTSKE  
 GSSSSVWFSPSPVLRAIAQREGIGLGNLOKIHAGTCKGGVTRQDL<sup>2</sup>EAIVISQOQVPIE  
 IFOGEVNRIPMSPLRAIASLSSKSDDEVLAVSDVVDVTDLKNLISGQRQRFLOTH  
 GVKLITSTVOCLOAOTLQFPLLSDGHTIYMKSVN<sup>2</sup>GVAVNLNKGCVYVPIVH  
 NCDRGLVSTAKALDOLSRRLNKLDFSEVDQSVYVNF<sup>2</sup>GMGTGALIGAPFIRIPEY  
 ALIGIGTIQRVVRDDDSLAKRWYVITLFDHKLVDGIGYSEFISLANKRLESVTM  
 G"  
 complement(9038. .10282)  
 /gene="gltT"  
 CDS complement(9038. .10282)  
 /gene="gltT"  
 /codon\_start=1  
 /transl\_table=11  
 /product="glutamate symport"  
 /protein\_id="BAA98734.1"  
 /db\_xref="GI:8978899"  
 /translation="MKLNKKIFGLVGVTLGLVLEDKAIFEFKIGDIFNL<sup>2</sup>SMVY  
 PLVCSMVLGIASI<sup>2</sup>DMKLKGRIGIKSVGLVLTALVIGLFCWAFISPGNGCQDA  
 QAOSMDSVATIDNKTAFLVSLIAQVPPVPSFAEGNLI<sup>2</sup>ILIFATLGLAGL  
 SGERGRVERIFDQSEIMLVNMTSFPAGVYGASMAWISGHNIGLVLQWLGKTL  
 IYLA<sup>2</sup>CLFHTNLVGLGRFCGKMSFSLMDPAISCAVSTASSATLIPVTRCVS  
 KNLGVSNHGVFLPGVAFVNNCTAIFQGNVAFIAQYNCPLSLSLULLLVYVITAF  
 SAVGSAGVGGGMITLGSVLASVGLPIGAILAGIDRLDIVGTIPNLLGDVAVYIT  
 VASGEGLSFIESIKQESVEIT"  
 complement(10282. .11379)  
 /gene="ycsA"  
 CDS complement(10282. .11379)  
 /gene="ycsA"  
 /codon\_start=1  
 /transl\_table=11  
 /product="AtPase"  
 /protein\_id="BAA98735.1"  
 /db\_xref="GI:8978900"  
 /translation="MKKRPSTLFLFYRRVTIAISLEGILGWGLGSLLSKVF<sup>2</sup>FAFLVA  
 CWNRFSTPYRARSTVSVGNVIRVYGGACKPTVLWLAELRLRGYSCGLVSRGYSQ  
 SDGQYKLLHVDKSVASVYSGDEPLLMAEKLPEGSVWVKDR<sup>2</sup>LSAARAEKFGCILL  
 SSROGLKLLHVDKSVASVYSGDEPLLMAEKLPEGSVWVKDR<sup>2</sup>LSAARAEKFGCILL  
 TVVKRYSNAPQIPVKPTIASVSVYTHNGIRIPKEALRELGVGFCGLGFPQGLNMLRE  
 EGHILGLVLLPDHAITTKKELNYFCQOMARQGGGLLCTERDSVKLPRUSGEVSLIP  
 IAKVEMRLSYNPDQ<sup>2</sup>LSLNLIEQHKHNGN"

D	b	41040	GCAAAATGTAACACAGTGTGTAATAGAGCCCTTGTGTACAAAGTAAATATCTCTGGTCTGAT	40981	LOCUS	AE001640	10422 bp	DNA	linear	BCT 01-DRC-2000
Q	y	1141	tggcttcacgtatgataaacctgtggagtagtactctatctcagtcgaataccctcggagacttg	1200	DEFINITION	Chlamydia pneumoniae section 56 of 103 of the complete genome.				
D	b	40980	TGGTCTTACGTATGTAACCTGTGGAGTAGTCTCTATCTCAGTATCGAATCTCTGGAGACTTG	40921	ACCESSION	AE001640.AE001363				
Q	y	1201	gtcttcacgtatgctggtatcccaagatacactccctctcgtggttacagtagtactcgaagct	1260	VERSION	AE001640.1	GI:4376845			
D	b	40920	GTCTCTCATGATGTCGATGATCCCAAGATACACTCCCTCTCGGTGTACAGTACTCGAAGCT	40861	KEYWORDS	Chlamydia pneumoniae CwL029.				
Q	y	1261	ctggttgagagatcgtcgtgataaaagtgtgttgcggtattaaagaataatgtgccaggga	1320	SOURCE	Chlamydia pneumoniae CwL029.				
D	b	40860	CCTGGTGGAGAGATCTGCTGTAATAAGTTGTTTGGCGTATTAAAGAAATGTGCCCAGGA	40801	ORGANISM	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia; Chlamydia pneumoniae CwL029.				
Q	y	1321	gaaacctccagtttaaaactgttagtgaaagctcaagttccctcgtggaagattccacaaataca	1380	REFERENCE	1 (bases 1 to 10422)				
D	b	40800	GAAACCTCCAGTTTAAACTGTGTAGTGAAGCTCAAGTTCCTGGGAAGATTTCACAAATCAA	40741	AUTHORS	Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R. W., Olinger, L., Grimwood, J., Davis, R. W. and Stephens, R. S.				
Q	y	1381	gttcagtaactagtagtcaactcggaacatgtacatcttcgcgagaaacacaaaca	1440	TITLE	Submitted (01-DRC-1998) Program in Infectious Diseases, University of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA				
D	b	40740	GTTCAGTAACTAGTGAAGTCTAACTCGGAACATGTACATCTTTCGCAGAAACAAACA	40681	JOURNAL	Location/Qualifiers				
Q	y	1441	cattggaaggtcttcagctacccatattcgttattagacacaaatgacctctctgt	1500	FEATURES	1..10422				
D	b	40680	CATTGGAAGGTCTTCAGCTACCCATATGCGTATTAGACACAAATGATCTTATCTGT	40621	source	/organism="Chlamydia pneumoniae CwL029"				
Q	y	1501	gtaggagaataactgctctatctcgtgttaactaaacccgtggtctcgtcgtgaagatact	1560	gene	/strain="CwL029"				
D	b	40620	GTAGGAAAATATCTGTCTATCTGTGTAACTAAACCGTGTCTCTGCTGAAGATACT	40561	CDS	/db_xref="taxon:115713"				
Q	y	1561	aacgtatctttaactctggaagttctcaaaagaacttcagcccaatagctcttcaggtcca	1620		84..374				
D	b	40560	AACGTATCTTTAATCTTGAAGTCTCAAAAGAATCTTCAGCCAAATAGCTTCTCAGGTCCA	40501		/gene="CpN0554"				
Q	y	1621	actaaagaacgatttcagtaataccgtgttttcgacgctttaccataactcagttct	1680		/gene="CpN0554"				
D	b	40500	ACTAAAGAACGATTTCAGGTAAATACCGTGTGTTTCGACGCTTACCTAAACTCGTCTCT	40441		/transl_table=1				
Q	y	1681	aaggaactgtagagttttctgttaactgaaaggtattgctcccgagagctcgcggc	1740		/product="Cp440 hypothetical protein"				
D	b	40440	AAGGAATCTGTAGAGTTTCTGTGTACCTTGAAGGTATTGCTCCCGGAGATGCTCGCGC	40381		/db_xref="GI:4376848"				
Q	y	1741	gaagctattcttctgtatcacactgactccaccagtagtaccagagcagatattccgctctcttaggat	1800		/translation="MYEGRMASPTPGOLHLOOKVESKAYDYSRLAMINATALLFFI				
D	b	40380	GAAGCTATTCTTCTGTATACACTGACTTCACCAGTATCAGACACAGAAAATACCCAC	40321		528..2474				
Q	y	1801	gtgtattaaatctcaaggaattatctcaagcagagcagatattccgctctcttaggat	1860	gene	/note="CpN0555"				
D	b	40320	GTGTATTAAATCTTAAGGAATTATCTTAAGCAGACCGATATTCCGCTCTCTTTAGGAT	40261	CDS	528..2474				
Q	y	1861	agctttcaagaagtagcgcgttagtagtaactcagtagtaaaagcgtttttttgttttata	1920		/gene="tsp"				
D	b	40260	AGCTTTCAAGAGAGTACCGCTTTAGTACCTTACGTACTTAAGCGGTTTTTTTGTTTTATA	40201		/codon_start=1				
Q	y	1921	agctttcaaatccaatgtagtagttctttaaatacaagatattatttaagttcttgaatac	1980		/transl_table=1				
D	b	40200	AGCTCTTCAATCCAATCGTAGTCTTCTTAATCAAGAGATATTATTAAAGTTCTGAAATC	40141		/product="15 kDa Cysteine-Rich Protein"				
Q	y	1981	ctaagatttattttaaaagccactcttttagtagtagtaattcaaatitttaattaaagct	2040		/protein_id="AAD18696.1"				
D	b	40140	CTAAGATTATTTTAAAGCCCATCTTTTAGTATGTAATTAATTTTAATTAATGT	40081		/db_xref="GI:4376848"				
Q	y	2041	tttctcagtagtaactcgtctctttaggaactacacacacacacacacacacacacacacac	2100		/translation="MSSNLHPVGGTGTGAAPESVLNIVEELAAAGSVTAGLOAITSS				
D	b	40080	TTTCTCAGTAGTAACCTGCTCTTTTAGGAACACTACACTAGGAGAACGATGTCATCAATAC	40021		PGMNLIGWAKTFIQPIRESKLFOSRACQITLLVLGILLVAGLACMFIHISOLGA				
Q	y	2101	tacatcccgta 2111			NAFWLIPAAIGLKLKLVLSLCPDEACTSEKLMVFOKWAGVLEDDDDGILLNNSKIP				
D	b	40020	TACATCCCGTA 40010			CHVTEGNTSRATTPVINDRGTPVLSPLVKIARY"				
					RESULT	4				
					AE001640/c					



Db 4569 GATCGCTGGTGCAGGAGATAAATGCTAAATTAATCTGATGGGTAAACCTCTTTAAAGAA 4510  
Qy 721 gttgtgtctcaagctgctactgtatgtgcttgcocagagctcgttcttactataa 780  
Db 4509 GGTGTGCTGCTTCACAGCTGCTACTGTATGTGCTTGCAGAGCTCCGTCTTATACATAA 4450  
Qy 781 tgcggtcaaccagccatttggattaaagcaagaagacctgactgtgcttgcctaaagatgc 840  
Db 4449 TCGCGTCAACAGCCATTGTATTAAAGCAAGAAGACCTGACTGTGCTTGAAGATGC 4390  
Qy 841 cctgtatgtcacaaatcaaatcaagtagtgacacagagatgctattgcccgttaacgttaact 900  
Db 4389 CCTGTATGCTACAAATCGAAGTAGTGAACACAGGATCTGCTATTGCCCCGTAACGTAACT 4330  
Qy 901 gtatagataactgctcccgatggctattctcatgcatctggttgcataagagttctctcttt 960  
Db 4329 GTAGATAATCCTGTTCCGATGGCTATTCTCATGCTATCTGGTCAAAAGAGTTCTCTCTTTT 4270  
Qy 961 aacttaggagacatgagacctggcgataaaaaggattattacagttgagttctgcccctcaa 1020  
Db 4269 AACTTAGGACATGAGACCTGSCGATAAAAAGGTATTACAGTTGAGTCTCTGCCCTCAA 4210  
Qy 1021 agaagaggtcaaatcaactgaacttactactgaacttactcgtcgtgagacacaaatgtct 1080  
Db 4209 AGAAGAGTCAAAATCACTAACGTTGCTACTGTACTGTACTTTCGCGTGGACACAAATGTTCT 4150  
Qy 1081 gcaaatgtaactacagttgttaattgagcccttgtgtacaaagtaaatatctctgtgctgat 1140  
Db 4149 GCAAAATGAATCACTACAGTTGTTAATGAGCCTTGTGTACAAAGTAATATCTCTGGTCTGAT 4090  
Qy 1141 tggcttcaactgtataaacctgtggagactctctactcagttacatcctcgtgagacttg 1200  
Db 4089 TGGTCTTACGTATGTAACCTGTGGAGTACTCTATCTCAGTATCGAATCTCTGGAGACTTG 4030  
Qy 1201 gtctctcatatgctgtatcaagatacaactccctctgtgtgttacaagtactcgaagct 1260  
Db 4029 GTTCTCTATGATGCTGTGATCCAGATACACTCCCTCTCTGTTGTACAGTACTCGAAGCT 3970  
Qy 1261 cctgtgtgagagatcctgctgtataaaagtgttggcgtatttaaagaatgtgccaggga 1320  
Db 3969 CTTGGTGGAGAGATCTGCTGTATAAAGTGTGTGGCGCTATTAAAGAAATGTGCCAGGA 3910  
Qy 1321 gaaacctccagtttaaacctgtagtgaagctcaagttccctgggaagattcacaaaacaa 1380  
Db 3909 GAAACCTCCAGTTTAAACTGTGTAGTGAAGCTCAAGTTCCTGTGGAAGATTACAAAACAA 3850  
Qy 1381 gtgcagtaactagtgagttcaactcgcggaacatgtacactcttgcgcgagaacaaacaa 1440  
Db 3849 GTTGCAGTAATAGTGAGTCTAACTCGGGAACATGTATCATCTTGGCGAGAAACAAACAA 3790  
Qy 1441 cattggaaaaggtcttcagctaccctatgtgctgattagacacaaaatgatccctatctgt 1500  
Db 3789 CATTTGAAAAGGCTTTCAGCTACCCATATGTGCGTATTAGACACAAATGATCTCTATCTGT 3730  
Qy 1501 gtaggagaaaactgtctatcgtatcctgttaactaacctgggttctgctggaagatact 1560  
Db 3729 GTAGGAAATAACTGTCTATCGTATCTGTGTAACCTGTTGCTGTGAAGATACT 3670  
Qy 1561 aagctactttaactcttgaagttctcaaaagaacttcagcaaatagcttcttcagggtcca 1620  
Db 3669 AACGTATCTTTAATCTGTGAAGTTCTCAAAAGAACCTTACGCCAATAGCTTCTTCAGGTCCA 3610  
Qy 1621 actaaaggaacgatttcaggtaaacctgtgttttcgcagcgtttacccctaaactcgttct 1680  
Db 3609 ACTAAAGGAACGATTTTCAGGTAATACCGTGTGTTTTCGACGCTTTTACCTTAAACTCGGTTCT 3550  
Qy 1681 aaggaactgtagagtttctgttaccttgaaggtattgtctcccgagatgctcgcggc 1740  
Db 3549 AAGGAATCTAGAGTTTCTGTACTTGAAGGTATTGCTTCCCGGAGATGCTCGCGGC 3490  
Qy 1741 gaagctattcttctctgatatacactgacttccacagttatcagacacagaaaataccccc 1800  
Db 3489 GAAGCTATTCTTCTTCTGATACACTGACTTTCACCAGTATCAGACACAGAAAAATACCCAC 3430

Qy 1801 gtcgtattaaattcttaagaataattatactaaagcagagcgatattccgctctgctttaggat 1860  
Db 3429 GTGTATTAATTTCTTAAGGAATATCTCTAAAGCAGAGCGATATTCGCGCTCTGCTTTAGGAT 3370  
Qy 1861 agctttcaaaagaagtcacgccttttagtactctacgtaaaagcgggttttttttttttata 1920  
Db 3369 AGCTTTCAAGAAGTACCGCTTTTAGTACCTTACGTACTAAAGCGGTTTTTTTCTTTATA 3310  
Qy 1921 agctcttcaatccaatcgtatagtttcttaatacaagataattttaagtttctcgaatc 1980  
Db 3309 AGCTCTTCAATCCAATCTGTAGAGTTCTTAAATCAAAAGATATTATTAAAGTTTCTGAATC 3250  
Qy 1981 ctaagatttatttttaaaagccatcttttttaggtatgataataaaatttttaattaaagt 2040  
Db 3249 CTAAGATTATTATTTAAAGCCCCATCTTTTAAAGTATGTAATTAATAATTTTAAATTAATGT 3190  
Qy 2041 ttctctagtgaacctgtcttttaggaactcacactagggaacgggtatgcatcacaatc 2100  
Db 3189 TTTCTCTAGTGAACCTGCTCTTTTAGGAACCTACACTAGGAGACGGTATGTCATCAATC 3130  
Qy 2101 tacatcccgta 2111  
Db 3129 TACATCCCGTA 3119  
RESULT 5  
AX349525  
LOCUS AX349525  
DEFINITION Sequence 48 from Patent WO0202606.  
ACCESSION AX349525  
VERSION AX349525.1 GI:18615369  
KEYWORDS Chlamydomophila pneumoniae.  
SOURCE Chlamydomophila pneumoniae.  
ORGANISM Chlamydomophila pneumoniae.  
REFERENCE 1 (sites)  
AUTHORS Ratti,G. and Grandi,G.  
TITLE Immunisation against Chlamydia pneumoniae  
JOURNAL Patent: WO 0202606-A 48 10-JAN-2002;  
Chiron S.p.A. (IT)  
FEATURES  
Location/Qualifiers  
source 1. 1671  
/organism="Chlamydomophila pneumoniae"  
/db\_xref="taxon:83558"  
BASE COUNT 488 a 342 c 370 g 471 t  
ORIGIN  
Query Match 79.2%; Score 1671; DB 6; Length 1671;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 139 atgtcacaactcatcagcagtagttacggtcccttgcctaacgagatcgcgagttgc 198  
Db 1 ATGTCCAAACTCATCAGACGAGTAGTTACGCTCTTGGCTTAAGAGTATGCGAGTTCG 60  
Qy 199 ttgcccagcgggggtatagagggcgcgtgtagcagagctctctgattactaagatcgct 258  
Db 61 TTGTCCAGCGGGGTATAGAGCCCGCTGTACAGAGTCTCTGTAGATTCGTCGCT 120  
Qy 259 agtgcggaacaaacccagcagcctgttctctatgacagcgaaggttagactgtccgt 318  
Db 121 AGTGGGGAACAAAGCCAGCACCTGTCTTATGACAGCGAAGAGGTAGACTTGTCCGT 180  
Qy 319 agaaataaacacacagtttgaacaaaagcgtgtgtcttctgtgataaagaattttat 378  
Db 181 AGAAATAAACACACAGTTGAACAAAAAGCCGTGCTGCTTTTGTGATAAAGAAATTTAT 240  
Qy 379 cccgtggaaggggagcagatgtaaccctgttagagctcagcaaggtcttctcagcaaga 438  
Db 241 CCCTGTGAAGAGGGAGCATGTCAACCTGTAGAGGCTCAGCAAGAGTCTTCTCTACGGAAGA 300



[illegible]



db	2597	GAGGAGGCGTATTCTTTCTTCAGACACTCTAACGGTACCTGTTGCTGATACAGAAACA	2656
QY	1796	ccacggtattaaatttaaggaattatctctaaagcagagcgatattccgctctgtgtt	1855
db	2657	CACATGTTTATTAAATCTTTAAATTTCTACTTTGAAAGATTGTAAACGAAAGGCGGT	2716
QY	1856	aggatagctttcaagaagtagcgcgttttagtacccttaagtaagcggtttt	1909
db	2717	CCTAGAAATTCCTAGGACGCGTCTTTTGTGTTCTCTTCATCGTCGATTTCT	2770
RESULT	7		
AF111201		1185 bp	DNA linear BCT 15-MAR-1999
LOCUS		Chlamydia pneumoniae strain AR388	outer membrane protein 2 (omp2)
DEFINITION		gene, partial cds.	
ACCESSION	AF111201	GI:4416362	
KEYWORDS		Chlamydia pneumoniae.	
SOURCE		Chlamydia pneumoniae	
ORGANISM		Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia	
REFERENCE		1 (bases 1 to 1185)	
AUTHORS		Kaltenboeck, B. and Gao, D.	
TITLE		PCR amplification and sequencing of the partial coding region of the outer membrane protein 2 Omp2 of Chlamydia	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1185)	
AUTHORS		Kaltenboeck, B. and Gao, D.	
TITLE		Direct Submission	
JOURNAL		Submitted (08-DEC-1998) Pathobiology, Auburn University, 270 Greene Hall, Auburn, AL 36849-5519, USA	
FEATURES		Location/Qualifiers	
source		1..1185	
		/organism="Chlamydia pneumoniae"	
		/strain="AR388"	
gene		/db_xref="taxon:83558"	
		<1..>1185	
CDS		/gene="omp2"	
		<1..>1185	
		/gene="omp2"	
		/codon_start=1	
		/transl_table=11	
		/product="outer membrane protein 2"	
		/protein_id="AAD20337.1"	
		/db_xref="GI:4416363"	
		/translation="IGKDCVDVITQQLPCAEFVSSDPETPTSDGLVWKIDRLG	
		AGDKKITVMVKPLKEGCGCTAATVCACPELRSYTKCGQPAICIKQEGDCAICLRCPV	
		CYKIEVNTGSAIARNVTVDNPVDPGYSHAGORVLSFNLGDMRPGDKKVFYVECPQ	
		RGQIINATVTVCGHKCSANVTTVNEPCQVNIISGADWSYCKPKYISVSNPGR	
		DLVLDVVQDILPSCGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVKVAQVPR	
		FTNOVATFSENGCTCTSCAETTHWKGLAATHMVCVLDNDPICVGENTVYRICVTR	
		GSADENVSILKFSKELQPIASSGPKGTISGNTVVFDALPKLGSKESVEFVSITLKG	
		IAP"	
BASE COUNT	347 a	238 c	255 g 345 t
ORIGIN			
Query Match		56.18;	Score 1185; DB 1; Length 1185;
Best Local Similarity		100.00;	Pred. No. 9.6e-257;
Matches 1185;		Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	541	ataggcaaaaagattggtgattgattacacacacagctacettgcgaagctgaa	600
db	1	ATAGGCAAAAAGATTGTTGATTTGATTACACACAGCTACCTTGGCAAGCTGAA	60
QY	601	ttcgttaagcagtgatccagaaacaaactcctcaagaatgaggaattagctggaatc	660
db	61	TTCGTAAGCAGTGATCCAGAAACAACTCCTACAAGTGATGGGAATTTAGCTGGAAATC	120
QY	661	gacgcgtgggtgcagagataaatacgaattactgtatgggttaaacctcttaaaaga	720
db	121	GATCGCTGGGTGCAGAGATAAATGCAAAATTTACTGTATGGTGAACACCTCTTAAAGA	180

QY	721	ggttgctgtccacagctgctactgtatgtgcttcccagagctccgttcttatactaaa	780
db	181	GGTTGCTGCTTCACAGCTGCTACTGTAATGCTTGCCACAGCTCCGTTCTTATACTAAA	240
QY	781	tgcggtcaaccagccatttatttaagcaagaagagacctgactgcttgccttaagatgc	840
db	241	TGCGGTCAACCGCAATTTGTTATTAAAGCAAGAGAGCTGACTGTGTTGCTGAAGATGC	300
QY	841	cctgtatgtctcaaaaatagaagtagtgaacacagagatctgctattgcccgttaagtaact	900
db	301	CCTGTATGCTACAAAATCGAAGTAGTAGAACACAGAGATCTGCTATTGCCGTAACGTAAC	360
QY	901	gtaataatcctgttcccgatggctattctcatcatcattggttgcgaagattctctttt	960
db	361	GTAGATAATCTTGTTCCCGATGGCTATTCTCATGCATCTGGTGAAGAGTTCCTCTCTTT	420
QY	961	aaacttagagacatgagacctggcgataaaaaggtatttcaagttgagttctgcctcaa	1020
db	421	AACTTAGGAGACATGAGACCTGGCGATGAATAAAGGTATTACAGTTGAGTTCTGCTCAA	480
QY	1021	agaagaggtcaaatcactaacgttgcctactgttaacttactgcggtggacacaaatgtct	1080
db	481	AGAAGAGGTCAATCACTAACGTTGCTACTGTAACCTTACTGCGGTGGACACAAATGTTCT	540
QY	1081	gcaaatgttaactacagttgttaagagccttgtgtacaaagtaaatatctctggtgctgat	1140
db	541	GCAAAATGTAACTACAGTTGTTAATGAGCCTGTGTACAAGTAATATCTCTGCTGCTGAT	600
QY	1141	tggtctacgtatgttaaacctgtggagctactctatctcagttacgtatcgcgaagacttg	1200
db	601	TGGTCTTACGTATGTAACCTGTGGAGTACTCTATCTCATGATCGAATCTCTGGAGACTTG	660
QY	1201	gttctcatgatgctgctgacgaagataacacctcctctggtgttacagttactcgaagct	1260
db	661	GTTCTTCATGATGCTGCTGATCCAAAGCTCAAGTCTCCCTCTGCTGTTACAGTACTCGAAGCT	720
QY	1261	cctgtgagagatctgctgtaataaagtgttggcgattattaaagaataatgcgcagga	1320
db	721	CCTGTGGAGAGACTGCTGTATAAAGTGTGTTGGGCTATTAAAGAAATGTGCCAGGA	780
QY	1321	gaacccctccagtttaaaactgtagtgaaagctcaagctcaagttcctggaagattcacaaatcaa	1380
db	781	GAACCCCTCCAGTTTAAACTTTGATGAAAGCTCAAGTCTCTGGAGATTTCACAAATCAA	840
QY	1381	gttcgttaactagtgagcttaactgcggaacatgtacatcttgcgcgaagaacaaca	1440
db	841	GTTGAGTAACCTAGTGTGAGTCTTAACCTGCGGAACATGTACATCTTTCGCGAACAACAACA	900
QY	1441	cattggaaggtcttgcaagctcccatatgtcggtattagacacaaatgactctctgt	1500
db	901	CATTGGAAGGCTTTCGAGCTACCCATATGTGCGTATTAGACACAAATGATCTCTCTGT	960
QY	1501	gtagggaataatactctctatcgatctgttaactaacccgtggttctgctggaagatact	1560
db	961	GTAGGAGAAATAACTGCTCTATCTGTATCTGTGTAACCTAACCGTGGTGTCTGCTGAAGATCT	1020
QY	1561	aacgtatctttaaactctgaagttctcaaaagaacttcagccaatagcttctcaggtcca	1620
db	1021	AACGTATCTTAACTTGAAGTCTCAAAAGAACTTCAGGCAATAGCTCTTCAGGTCCA	1080
QY	1621	actaaaggaacgatttcagtaataacggttttctgcagcgttttactaaactcgtgtct	1680
db	1081	ACTAAAGGAACGATTTTCAGGTAATACCGTTGTTTCGACGCTTTACCTAAACTCGGTCT	1140
QY	1681	aagaaatctgtagagtttctgttaccttgaaaggtattgtctccc	1725
db	1141	AAGAAATCTGTAGATTTTCTGTTACCTTGAAGGTATTGCTCCC	1185
RESULT	8		
CPSCROMP		2478 bp	DNA linear
LOCUS		CPSCROMP	



```

1107 TTGAGTTGTGACAGAGTATCTCTGCGACCAACACCACTCGGATAGCAAAATTAATCTGGA 1166
QY 656 aaatcgtcgcgtgggtgcaggagataaaatgcaaaattactgtatgggtgaaacacctta 715
Db 1167 CAATGTATCGCTAGGTCAAGGTGCAAAATGCAAAATACCGTTTGGGTGAAACCTCTTA 1226
QY 716 aagaagttgtctctcagagctactgtatgtctgtgcccagagctcgtctcttata 775
Db 1227 AAGAAGGTGTGCTTACCGCGGCTACTGTATGTGCTTGCCGCAACTTCGCTCTTATA 1286
QY 776 ctataatcggttaacacagccatttgtatttaagaagaagacactgactgtcttgccata 835
Db 1287 CCAATGCGGACACCAACAGCTATTTGTTATTAAACAAGAGTCCCTGAATGCCGCTTGCCTAC 1346
QY 836 gatccctgtatgctacaaaatcgaagttagtgaacacagagatctgtctattgcccgtaacg 895
Db 1347 GTTGCCAGTGTGTACAAAATGCAAGTTTGAACACAGAGTTCTGTATAGCCCGCTAATG 1406
QY 895 taactgtagataaactcctgtcccgatgctatctctcatgcatctgttcaaaagagttctct 955
Db 1407 TTGTCGTGATACCCAGTTCGCCGATGGCTATACCTCATGCTTCAGGACACGCGTCTCTT 1466
QY 956 ctitttaacttaggagacatgagaccctggcgataaaaaaggtatttaccagttgagttctgcc 1015
Db 1467 CTTTAACTTAGGAGATATCGCTCGGGATTTCTAATGCTCTGTGTGGAGTTTGGCC 1526
QY 1016 ctcaaaagaggtcaaatcactaacgttgcctactgtacttactgcgtggtggacacaaat 1075
Db 1527 CACAAAAAAGAGGAAAGTTACTAAGCTGGCTACTGTATCTTACTTGCAGGAGACATAAT 1586
QY 1076 gtctgcaaatgaactcacagttgttaagtgcctgtgtacaaagataaatactctggtg 1135
Db 1587 GTTCTGCTAAGCTAAGTACTGTAGTAAACGAACCACTGCTCAACAGTAATAATCTCTGGAG 1646
QY 1136 ctgattggtcttaogtagtaaacctgtgagtaactctctcagatcgaatactgagag 1195
Db 1647 CTGACTGTGCTTATGTATGTAAGCTGTAGTAATCACTATCTGTTATCTTAACCCAGAG 1706
QY 1196 acttggttcttcattgctgatacgaagatacacctcctcttctggtgttacagtaactcg 1255
Db 1707 ATCTTAAACTTTAGGATGTAGTTATAGAAGATACCCGACCTTCAGGAGCAACAATTTTAG 1766
QY 1256 aagctccgtggtgagagatcgtgtaataaagttgttggcgattataaagaagattgccc 1315
Db 1767 AAGCTGCTGGAGCTGAATCTGCTGTAACAAAGCTGTATGTGTGATCAAGAAATGTGCC 1836
QY 1316 caggagaacccctccaagtttaactctgtagtgaagctcaagtctcctggaagattccaa 1375
Db 1827 CAGGAGAGACTCTCAATTTAAAGTTGTGTGCTAAAGCAACAAAGCCAGTAAATTCACA 1886
QY 1376 atcaagttgcagtaactagtagtcttaactgoggaacatgtacattcttgcgcaaaaaca 1435
Db 1887 ATCAAGTTGTTGTCAAAACTAATCCGATTTGTGGAACCTGTACTTCTTCCGCAAGATTA 1946
QY 1436 caacacattgaaaggtcttgacagtlaccatattgctgtatttagacacaatgatccta 1495
Db 1947 CAACACATTGAAAGGCTTTCAGCTAGCGATATGTGCGTAATCGATACCAATGATCCTA 2006
QY 1496 tctgttagagaataactctctatcgtatctgtgttaactaacctggttctactaaag 1555
Db 2007 TTTGCGTAGAGAAATACCTGTATACCGTATTTGTGTACCAACACCGTGTCTCGAGAG 2066
QY 1556 atactaagtagtctttaaattgaagttctcaaaagaactcagccaatagttcttctcaag 1615
Db 2067 ATACGAACGCTCTGCTTAATCCTTAAGTTTCTTAAGAACTACAAACCCGCTTCTTCTTAG 2126
QY 1616 gtccaaactaaaggagacttccagtaataaccggtgttttgcagcgtcttaccataaactcg 1675
Db 2127 GTCCAAACAAAGGAACCAATTACAGGGAATACAGTAGTATTCGATGCTCTACTCTAAATTA 2186
QY 1676 gtctaaaggaactgtagaatttctgtacctgaaagtagtctcccgagagactc 1735
Db 2187 GTTCTAAGGAATCTGTAGAGTTTCTGTGAACATTAAGAGGATTTGCTCCAGGAGATGCTC 2246
```

QY 1736 gcgcgcaagctattcttctgtatcacactgactccagctaccagatcacacacaaaata 1795  
Db 2247 GAGGAGAGGCTATTCTTCTTCTAGACACTTAAACGCTACCTGTTGCTGATACAGAAAACA 2306  
QY 1796 ccacgctgtataaattcagaagaattatctctaaagcagcagatattccgctctgcttt 1855  
Db 2307 CACATGTTTATTAATCTTTTAAATTTCTTAAAGATTTGTAACGAAAAGGCGGT 2366  
QY 1856 aggtactgtctaaagaagctacgcgttttagctactacgtactactaaagcggttt 1909  
Db 2367 CTAAGAAATCTTAGGAGCGCTTTTGTGTTCTCTTATCATCTCAGATTTTCT 2420

RESULT 9  
CPU41759 8950 bp DNA linear BCT 07-MAY-1999  
LOCUS  
DEFINITION Chlamydomophila caviae strain Guinea Pig Inclusion Conjunctivitis  
(GPIC) RecJ recombination protein gene, partial cds; glutamyl-tRNA  
synthetase homolog, outer membrane protein 3 (omp3), outer membrane  
protein 2 (omp2), and hypothetical sulfur-rich protein (srp) genes,  
complete cds; and unknown genes.  
ACCESSION U41759  
VERSION U41759.1 GI:1783376  
KEYWORDS cysteine rich protein; outer membrane protein; glutamyl-tRNA  
synthetase.  
SOURCE Chlamydomophila caviae.  
ORGANISM Chlamydomophila caviae  
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
REFERENCE 1 (bases 1 to 6000)  
AUTHORS Hsia, R.C. and Bavoil, P.M.  
TITLE Homologs of Escherichia coli recJ, gitx and of a putative 'early'  
gene of avian Chlamydia psittaci are located upstream of the 'late'  
omp2 locus of Chlamydia psittaci strain guinea pig inclusion  
conjunctivitis  
JOURNAL Gene 176 (1-2), 163-169 (1996)  
MEDLINE 97075925  
REFERENCE 2 (bases 6001 to 8950)  
AUTHORS Hsia, R.C. and Bavoil, P.M.  
TITLE Sequence analysis of the omp2 region of Chlamydia psittaci strain  
GPIC: structural and functional implications  
JOURNAL Gene 176 (1-2), 155-162 (1996)  
MEDLINE 97075924  
REFERENCE 3 (bases 1 to 8950)  
AUTHORS Hsia, R.-C. and Bavoil, P. M.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-1995) Ru-ching Hsia, Microbiology & Immunology,  
University of Rochester, 601 Elmwood Avenue, Box 672, Rochester, NY  
14642, USA  
FEATURES  
source  
1. .8950  
/organism="Chlamydomophila caviae"  
/strain="Guinea Pig Inclusion Conjunctivitis (GPIC)"  
/specific\_host="guinea pig"  
/db\_xref="taxon:83557"  
/notes="submitted as Chlamydia psittaci"  
CDS  
<1..809  
/note="ORF1: homolog of Escherichia coli recJ"  
/codon\_start=3  
/transl\_table=11  
/product="RecJ recombination protein"  
/protein\_id="AAB4138.1"  
/db\_xref="GI:1783377"  
SSNDHARVITIISAKLAKYKPKVAIISNGGKSLRTIGSFLLGILQKCSLFL  
ISYGGHFAAGIIINDEOYAEAFKFKIHLVNSLSKDKKAVLTLPDAQADFIDHDL  
LSSIDLFPFGKGNPMFPVYITVHQVRYPKLLNHLKLVNHLNGLERLEGAFGLGDR  
IEALKANNKPLELAYTPRLSOSNGSVIHLVRFDFILPLNYKDSQPKF"  
1250..2404  
/note="ORF2"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"

```
/protein_id="AAB41139.1"
/db_xref="GI:1783378"
/translation="MFSVQSSVIGNSLITMNPQETLIFSSFAAQAHRRTLYPVI
YKLDVICELIKCIPRIILIFPGLLVGLKICQNCGLLPAAGGLETFEOLCFKAKILOE
SPRCYVDWLESYASSYKRIPIQCDOLLIDAISIOPPDAEKDRMLYSLGNSCEPFT
RAMVSRNDWILNVAQKARAVVYFNYPGYMHSKGPITRESLCKAYQACVYHLRHPD
GPKAGDIIAGSLGLYQALGSKREYVDSGDVDMFVIKDRGRPSVSAKIASOWLGRV
AGWTATQWEINSKISSELSVCFPIGHEDTQALIGDGMFNDRNCPAEPFLDPNA
PFPFGKKIPVGEFLLOHKGEDETTONVVKHIIIDHYPSPETPDGGE"
3005..3550
/notes="ORF4; homolog of EUO of Chlamydia psittaci 6BC"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAB41140.1"
/db_xref="GI:1783379"
/translation="MECIQHSCFDLDDRENAQOLEVOEGTEMVSTOAAKLNHVTRQ
AIVYAKOKKLAASKTTRWEIDLKDEYKRNYSRKKSLYOGELLFNDMDKGFYSVNO
VADMLGIPQVKYVYATRTGMRGERKGAANVISQSELDIRYKSEYLNKQIVKKKEVAV
EHTATPADAVESGTALFEND"
3822..5339
/notes="ORF5; homolog of glx of prokaryotes"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA synthetase homolog"
/protein_id="AAB41141.1"
/db_xref="GI:1783380"
/translation="MAWENVRVRVAPSGTPDPHVTAYMALFNEIPAKRFNGKMILRI
EDTDOTSRDDYKEKNIFISALQWGIQWDEPDIGGPHGYRQSERTEIYREVAELLLK
TDYAKCFATPEKLEMRVATLILRGYDRYRVLSPSEIEARTQEGQPTIIRLVK
PLTGEVCLEDYCKGRVFPWADVDQVLMKSDGEPYHFANVVDHLMGITHLVGRV
WLSSTPHLLLYAFRGWPPIFLHMLPLLPDGTGKLSKRNPTSIIFYRDACYKEAF
NPLFLMGYSMEGDEYVLSLEKLANEDPKICKSGAVEDVRKLDMWNKLYLHNSGP
ENLLALRWLNDRELLKILPLCSRMATLASPVGLSEPFESVLPYKSEELLPAAI
SQEKAILEYSYVYKLEKTDLWVQDFVLASKLSEAFQVHHKKVYIPLLYVAIIGKK
QGPLEFDSMELUGKEPTRMWVHAQNLGGVPKAIQIADIKLKEENFNKILFE"
5992..6258
/gene="omp3"
/gene="omp3"
/codon_start=1
/transl_table=11
/product="outer membrane protein 3"
/protein_id="AAB41142.1"
/db_xref="GI:1783381"
/translation="MKKAILLAACVGVGLSSCCRTVDCCFEDPCAPKSCNPQVFK
KDGECSPCGSYTPSCSKPCGSDCAKVGQPAKGTSPDGRCKQ"
6423..8099
/gene="omp2"
/gene="omp2"
/notes="developmentally-regulated; cysteine-rich"
/codon_start=1
/transl_table=11
/product="outer membrane protein 2"
/protein_id="AAB41143.1"
/db_xref="GI:1783382"
/translation="MSKLFRRVTVLALTSMASFASGKTEAAAEASLATRIFAEN
SDNISQSRKAKVRGRNKKYQOKNNGPYCKEFPYCDGSCQSSVDTKQPCYVK
MVCYRSDSNVIEISOAPEVATVGSVPYIELAVGKDCVNVVITQOLPCYVEFIS
DPATPTSDSKLIWTDILGGERCKITVMWPLKKECCFTAATVACAPCLSYTKCG
QPAICIKQSGPCACLRCPVYKILEVNTGSAIARGVVDNVPDGYSPRSQORVLSF
NLDMPGDKCTVFECQKRGKITNAVIVSGGHKCSANVTIVINPCQVNIISG
NMCEYLQKPEVETIVVSNPDLKIDVDVIEDLIPSGITLLEAPGAELCCNKAWCIE
MDSGETLQKPVAKAQTPKGTQNVKNTSDGSCFCAEVTTHWKGLAAITHMVID
TNDPICVGNENTYRIVCTNRGSAEDTNVSLILFKSELOPVSSSGPKGTITGTWVF
DALPKLGSKESEYFEFVTLKGVPADGARGEAILSSDILTVPVADTENTHY"
8310..8792
/gene="srp"
/gene="srp"
/codon_start=1
/transl_table=11
```

```
/product="hypothetical sulfur-rich protein"
/protein_id="AAB41144.1"
/db_xref="GI:1783383"
/translation="MLTGVNSESSEGVIDILKPLCLDDVMKNETVOVTLVNSLVGWCKAH
IVDPIKSKIVQSPAFQIOTMVVLGVILLTAGLALFVLQGLGKNAFLELIPAVIGLV
KLLTSPNFKPCTPEAKWLCKRLALTATEDILLDDQIQNSNTLFTTSSSDVNTATOS"
BASE COUNT 2724 a 1612 c 1858 g 2756 t
ORIGIN
Query Match 54.9%; Score 1158.8; DB 1; Length 8950;
Best Local Similarity 78.4%; Pred. No. 6.2e-251;
Matches 1440; Conservative 0; Mismatches 387; Indels 9; Gaps 4;
Qy 1 ttgatcagtagtaggagatgaataattctgaacttctgaacttaattca-gataataaac 59
Db 6286 TTGGTTAGGTAGTAGGAGGTAAATATTTCCTGCTTAACTACCTTTATGAAAAATAACT 6345
Qy 60 caaatgttgagggttaagagtttacaataattctaccgatggcagaagaaaaataaa 119
Db 6346 TAAATGTTAGGTAGTAGATC--CATCATTTCACCCGATGCGAAGAAAAATAACR 6403
Qy 120 catgcagtaggagatccctatgtccaaactcatcagacgagtagttacgtcttcgcgt 179
Db 6404 CACGCGATAGGAGATCCCTATGTCCTCAAACTCTTCAGACGAGTAGTTACGGTCTCCGCT 6463
Qy 180 aacgagtagggcaggttgctttggcagcgggggtatagagccgcgtatagcagagtcct 239
Db 6464 AACTAGTAGGGCAGTTTATTTGCCAGCGCAAGATAGAGGCCGCTGCTGCACAGTCTCT 6523
Qy 240 gattactaagatcgtcgtcagtcagtcagaaa---caaacgcagcactgttctcagacgc 296
Db 6524 TGTACAAAGATTATTGCTAGTGTCTGAAACTCAGATGACANATTATTCGCAGAGAAAGC 6583
Qy 297 gaagaggttagacttgcgtccgtgaagataaacaacacaggtgaacacaaaaacgcgtgctc 356
Db 6584 CAAGAAGTGTAGATTGGTCGTAATAATAATCAAAAACAAGACAAAAAATAATGGCC 6643
Qy 357 ttttgtataaagaattttatccctgtgaagaggagcagatgtca---acctgtagaagc 413
Db 6644 TTACTGTGATAAAGAAATTTTATCCGTGTCAAGATGGCTCATGCCAATCATCAGTAGATAC 6703
Qy 414 tcagcaagagctctgctacggaagattgtattctgtataaagataaacgagtagttgcaact 473
Db 6704 TAAGCAAGAACCTTGTCTCGCAAAATGCTATTGCTGCTAGCTGATGATAGTAACGT 6763
Qy 474 agaaatttcaggtcgttcagaaatcagctactgttaggactctctaccatttgaagt 533
Db 6764 TGAATTAGCCACAGCTGTACCTTGAATATGCCAACAGTAGGATCTCTTATCCYATTGAAT 6823
Qy 534 ccttgctataggcaaaaaagattgctgtgattgattacacacagctactctgcga 593
Db 6824 TCTGTGCTGTAAGGTAAAGAGATTGCGTTAAATGTTGTTATCATCAACAGCTTCTCTTGA 6883
Qy 594 agctgaattcgttaagcagtgatccagagaacaactcctcacaagtgatgggaattagctcg 653
Db 6884 AGTTGAATTCATAAGCAGCGATCCTGCACAAACACTTACTTCCGATAGTAATAATTTG 6943
Qy 654 gaaattcagctcctcgttgtaggagataaataactactgtatgggttaaacctct 713
Db 6944 GACAATTGATCGCTTAGGTCAAGGTGAAGATGCAAAATTTACTGTTTGGGTAAACCTCT 7003
Qy 714 taagaaggttgctgctcagcagctgctactgtagtgcgtgcccagagctcctgctcta 773
Db 7004 TAAAGAAGGTGCTGCTGCACAGCTGTACCGTATGCTGTGCCAGAACCTTCCGCTCTTA 7063
Qy 774 tactaatgcgggtcaaccagcatttgtattaaagaagaaggacctgactgtgctgcct 833
Db 7064 TACCAATCGGACAACCTGCTTATTGTCATTAAAGCAAGAGGCCCTGAGTGCCTTGTGCTT 7123
Qy 834 aagatgcctctgtagctcaaaatcgaagtagtgaacacaggaatcgtctattcccgtaa 893
```

Tue May 28 08:55:07 2002

AF240773  
AF240773.1 GI:12619851  
Chlamydomophila abortus.  
Chlamydomophila abortus  
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
1 (bases 1 to 3186)  
Bush, R.M. and Everett, K.D.  
Molecular evolution of the Chlamydiaceae  
Int. J. Syst. Evol. Microbiol. 51 (Pt 1), 203-220 (2001)  
21078680  
11211261  
2 (bases 1 to 3186)  
Everett, K.D.E.  
Direct Submission  
Submitted (01-MAR-2000) Medical Microbiology and Parasitology,  
University of Georgia, College of Veterinary Medicine, Athens, GA  
30602, USA  
FEATURES  
Location/Qualifiers  
1..3186  
/organism="Chlamydomophila abortus"  
/strain="EBA"  
/specific\_host="Bos taurus"  
/db\_xref="taxon:8355"  
/note="isolated from cattle"  
100..363  
/gene="omlA"  
100..363  
/gene="omlA"  
/codon\_start=1  
/transl\_table=11  
/product="small cysteine-rich outer membrane lipoprotein"  
/protein\_id="AAG60549.1"  
/db\_xref="GI:12619852"  
/translation="MKKAVLLATVFCGALGLTSCRIVDCCFEDPCAPKPCNPGCNKK  
DKGSPGCTYPTSCTKPCGSECSGVOGQAKGCTSLDGRCKQ"  
529..2202  
/gene="cmcB"  
529..2202  
/gene="cmcB"  
/note="envelope protein"  
/codon\_start=1  
/transl\_table=11  
/product="60 kDa cysteine-rich membrane complex protein"  
/protein\_id="AAG60550.1"  
/db\_xref="GI:12619853"  
/translation="MSKLIIRVVTVLALTSMASFSAGKIEAAAAESLATRIASTEN  
SDNVLQATAKKVRGRNKNQEQKHTGAFCKEFYCEGCGOCQVDTTQESYGM  
YCVRNDDCNVEISQAPYAYVGSYPYPIELIAVGKKDCNVVITQQLPCEVEFYSSD  
PATPTSDSKLIWTIDLGGEKCKITVWVKPLKGCCTAATVCAACPLRSYTKGQ  
PATIKQEGPECACLRPCYKIEVCTGCTGALARNVVDNVPDGVTHASGQVLSFN  
LGMRPDSDKSCFVEFCPQKRGKITNVATSYCGGHKCSANVTYVNEPCVQVNSGA  
DMSYVCKPVEYITVNSLDKLYDVVVEDTVPSGATILEAEGAEICCNKAWNICEM  
CPQETIQKVVAKAOSPGKFTNOVVKTNSDGTCTSCAETTHMKGLAATMCVVDIT  
NDPICVGVNTYRICVTNRGSADTNVSLILKPSKELQPVSSGPKGTITHTVTVFD  
ALPKLGSKESEFVSYLKGIAPDARGEAILSSDTLTVPVADTENTHYV"  
2413..2817  
/gene="srp"  
2413..2817  
/gene="srp"  
/codon\_start=1  
/transl\_table=11  
/evidence="not experimental"  
/product="sulfur-rich protein"  
/protein\_id="AAG60551.1"  
/db\_xref="GI:12619854"  
/translation="MAGSTNSVGNDSITSLIQGLDQVIODEGVQVTLNLSILGWCRI  
HILNPVSKSKIVKRAFOITMIVIGIILLIAGLALTFLVQGLGNNAFLFLIPAVIGL  
VKLATSVFMEKPTPEKWLCKRLLOQLKIF"  
968 a 599 c 654 g 965 t

7124 AC GTT GCC AGT TTT GTT TAC AAA ATC GAA GTT ACC AAC CAC AGG TCT GCT ATAT ACC CGC GTG 7183  
894 CGT AACT GTAG AATA CTCTT CCAAT GCT ATCT CCA TCA TCA TCT GGT CAA AGA GTT CT 953  
7184 TGT TGT GGT TGA TAC CCA GTT ACC GAT GCG GTT GGT ATT CTT CAC CTT CAG GAA CCG CTT CT 7243  
954 CTCTTT AACT TAG GAC ATG GAC CTT GCG GAT AATA AAG GTT ATT TAC AAT TGA GTT CTG 1013  
7244 TTC CTT TAA CTT AGA GAT ATG CTT CCG TGG TGA TACA AAA TGC TTT ACT GTAG A GTT TTG 7303  
1014 CCCTCAA AAG AAG GTT CAA TCA AACT GCT GCT ACT GCT AACT TACT GCT GGT GAG ACAA 1073  
7304 CCCTCAA AAG AAG AAAAA TACC ACG GTT GCT ACT GCT ATCT TACT TACT GCG GTG GAC ATAA 7363  
1074 ATG TCT CCA AAT GAA CTA CAG TTT GTT AAT GAG CTT GTT GTC AAG AATA AAT ATCT CTT G 1133  
7364 GTG TTT CCG CTA ACG TTA CTT ACT TGT AAT CAA GAG CTT TGC GTT ACAA G TAA ATAT CTT CCG 7423  
1134 TGT GAT GTT GCT TCA CTA TGA AACT GGT GAG TACT CTT CTT CAG TAT CTT CAG TAT CTT C 1193  
7424 AAT GACT GGT CTT ACG TAT GAA CCG TGT AGA GTT ACC AAT TGT TGT ATCT TAA CCG CAG 7483  
1194 AAG TGT TCT CTA GAT GCT GAT CCA AAG AAT TTA TCC CCA TCC CTT CTT GGT TTA CAG TAT 1253  
7484 GGA TCT TAA CTT TAG AAT GTT GTT TAT AAG AAT TTA TCC CCA TCC CTT GGT TTA CAG TAT 7543  
1254 CGA GCT CCG TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG AAT 1313  
7544 AGA GCT CCG TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG AAT 7603  
1314 CCAG AAG AACT CCG TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1373  
7604 TCA GAG AAG AACT CTT CAA TTT AAG TTT GGT TTA AAG TTT GGT TTA AAG TTT GGT TTA A 7663  
1374 AAT CAA GTT TCC CTA AAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1433  
7664 AAT CAA GTT TCC CTA AAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 7723  
1434 AACA CCA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG AAT 1493  
7724 TACA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 7783  
1494 TAT CTT GTAG AAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG AAT 1553  
7784 TAT TTT GCG TAG TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 7843  
1554 AGA TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1613  
7844 AGA TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 7903  
1614 AGG TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1673  
7904 GGT TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 7963  
1674 CGG TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1733  
7964 AGG TCA CTA CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 8023  
1734 TCG GCG GAG CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 1793  
8024 CCG GCG GAG AAG CTA TGT GAG AAT GCT GAT AATA AAG TTT GGT GGT TTA AAG AAT TTA AAG 8083  
1794 TACC ACG TGT AATA AAT CTA AAG AAT TAT CTA A 1829  
8084 CAC ACG TGT AATA AAT CTT TTT AAA TTT CTT TAAA A 8119  
RESULT 10  
AF240773 3186 bp DNA linear BCT 07-MAR-2001  
LOCUS Chlamydomophila abortus EBA small cysteine-rich outer membrane  
DEFINITION lipoprotein (omlA), 60 kDa cysteine-rich membrane complex protein  
(cmcB), and sulfur-rich protein (srp) genes, complete cds.







CPU76761  
 LOCUS Chlamydia pecorum cystein-rich outer membrane protein omp-2 gene, BCT 26-JAN-1999  
 DEFINITION partial cds.  
 ACCESSION U76761  
 VERSION U76761.1 GI:4098316  
 KEYWORDS  
 SOURCE Chlamydia pecorum.  
 ORGANISM Chlamydia pecorum.  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
 REFERENCE 1 (bases 1 to 1604)  
 AUTHORS Sheehy, N., Markey, B. and Quinn, P. J.  
 TITLE Sequence analysis of C. psittaci and C. pecorum 60 kDa genes  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1604)  
 AUTHORS Sheehy, N., Markey, B. and Quinn, P. J.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-OCT-1996) Veterinary Microbiology and Parasitology,  
 University College Dublin, Ballsbridge, Dublin, Dublin 4, Ireland  
 FEATURES  
 CDS  
 1..1604  
 /organism="Chlamydia pecorum"  
 /strain="W73"  
 /specific\_host="Ovis sp."  
 /db\_xref="taxon:85991"  
 <1..>1604  
 /note="60 kDa protein"  
 /codon\_start=3  
 /trans\_table=11  
 /product="cystein-rich outer membrane protein omp-2"  
 /protein\_id="AA009598.1"  
 /db\_xref="GI:4098317"  
 /translation="TVLALSTMTSSFASGIEFAVAESLMTKIIANAESKRAADHKKE  
 AKHNRRKKSHATKHSNPNCCCKPEKKNQVQESQCFKMTYSIKVNDCC  
 NVLSQSPYATVSGSPYPIELALGKKDCNVNVIQQLPCEAEFVSDDPETPTADG  
 KLMMKIDRLQGGERCITVVMKPIKEGCFPTAATACGPEIRSTKCGQAICIKQEG  
 PDCALCPVRYKIEVTSKSAIARNVVDNVPDGFSHASQSVLSEGLMDHMPGES  
 KLVCEFCPKQKQVNVNATVYCGGSHKCSANMVIINPCVNVITISGVDPWVCKPK  
 EYLTGKQKQVNVNATVYCGGSHKCSANMVIINPCVNVITISGVDPWVCKPK  
 VLVKAQVPGQITQSVVNSQSGVCTSCADATVWKGLATHMCDVTDNDPCVGEN  
 VYRICITNRSAGDNTVSLILKFSKELQPISSSGPTKGTISGNTVIFDALPKLGSKE  
 SVESFVTLKGIAPCDARGEAVLFSDT"  
 BASE COUNT 498 a 326 c 345 g 435 t  
 ORIGIN  
 Query Match 44.88; Score 944.8; DB 1; Length 1604;  
 Best Local Similarity 74.88; Pred. NO. 1.3e-202;  
 Matches 1199; Conservative 0; Mismatches 402; Indels 3; Gaps 1;  
 QY 164 ttacgggtccttcgcgaacgagtagtcgtagtgcttgcgcagcggggatagagccg 223  
 Db 1 TTACGGGTCTTTCGCGTAACCTAGTAGTACGAGTTCCTTGCACGCGGGGAATAGAGCCG 60  
 QY 224 ctgtacgagctctctgattactaagatgcgtcgtagtcggaacaacagccagcacctg 283  
 Db 61 CTGTACGAGGCTCTGATGACTAAAATTATTGCAAAATGAGAACTTAAGCGTGCAGCAG 120  
 QY 284 ttccctacacgacgaagaaggttagactcttcgtagaataaacaacaccagttgaacaa 343  
 Db 121 ATCATAAAAAAGACCAAGCATAACTCGTAAACAAAAACATTCAAAAGCTACAAAACACA 180  
 QY 344 aaagccg--tggtgcttttggatgataaagaatttatccctgtgaagaggagcagatgc 400  
 Db 181 ACAGCAGATCTAACCCITCAATGTGTGTGATTAAGAGATTCCCTTGTGAAGAGAAGAATATGCC 240  
 QY 401 aacctgtagaggtcagcagaggtcttcgtacggaagattgtattctgtaaagtaaacg 460  
 Db 241 AACAAAGTAGAGAGCGACGAAGAATCTTGCCTTCGGTAAAGATGATTATTCATAAAAGTAATG 300  
 QY 461 atgattgaacgtgagaatttgcgcagtcctgtccagaataacgcactctgttagtatctct 520  
 Db 301 ACGATTGCAATGTAGCACTAAGTCAATCTGCTGCAATATGCAACAGTCGATCTCCAT 360

Query Match 42.1%; Score 889.6; DB 1; Length 2390;  
Best Local Similarity 71.0%; Pred. No. 3.4e-190;  
Matches 1249; Conservative 0; Mismatches 479; Indels 32; Gaps 4;

```
QY 1137 tgattgcttcacgtatgtataaactgtggagtaactctctctcagtatcgaaactctggaga 1196
|||||
Db 1607 AGATTGCTCTTATGTTTAAAGCCTGTAGAATATGCTGATCTCCGTTTCCAATCTCGAGA 1666
|||||
QY 1197 ctgtgtcttcacgtatgtatcccaagatacacactccctctctgtgttactacgtactcga 1256
|||||
Db 1667 TCTTGTGTTGCGAGATGTCGTGCTGAAGACACTCTTTCTCCCGAGTCACAGTCTCTGA 1726
|||||
QY 1257 agctccctggggagacatctgctataaagaatttggcggtatttaaaagaatgtgcc 1316
|||||
Db 1727 AGCTGCAGGAGCTCAAAATTCCTGTATTAAGTAGTTTGGACTGTGAAGAACTGAATCC 1786
|||||
QY 1317 agggagaacccctccagctttaaactgtagtgaagctcaagtctcctggaaagattccaaaa 1376
|||||
Db 1787 TGGAGAGTCTCTACAGTATAAAGTTCTAGTAAGAGACACAAACTCTCTGACAAATTCACAA 1846
|||||
QY 1377 tcaagttgcagtaactagctactaaactgcggaacatgtacatcttgcgcagaaacaac 1436
|||||
Db 1847 TAATGTTGTGTGAAGAGCTGCTCTGACTGTGGTACTTGTACTTCTTGCAGAACGAC 1906
|||||
QY 1437 aacacattgaaaggctcttcagctaccatattgtcgctattagacacaaatgatctctat 1496
|||||
Db 1907 AACTTACTGAAAGGAGTTGCTGCTACTCATATGTCGTAGTACTTGTGACCTGT 1966
|||||
QY 1497 ctgtgtaggagaaaaactgtctatctgtctgtgttaactaaacgtgtgtctgtgaaga 1556
|||||
Db 1967 TTGTGTAGGAGAAATACTGTTTACCGTATTGTTGTCACCAACAGAGGTTCTGCAGAAGA 2026
|||||
QY 1557 tactaactgtattcttttaattgttaagttctctaaagttctcaaaagaacttcagccaatagctcttcagg 1616
|||||
Db 2027 TACAATATGTTTCTTAATGCTTAAATCTCTTAAGAACTGCAACCTGTATCTCTCTGG 2086
|||||
QY 1617 tccaactaaaggagacattctcaggtaaacctgtttcttcagctcttaccctaaactcgg 1676
|||||
Db 2087 ACCAACTAAAGGAACGATTACAGGCAATACAGTAGTATTCGATTCTGTTACTAGATTAG 2146
|||||
QY 1677 ttctaaagaaactgtagattttcttctaccttaaaagttattctcccggaatctcgt 1736
|||||
Db 2147 TTCTAAGAAACGTAGAGTTTCTGTAACATTGAACAGCATATCAGCTGGAATGCTCG 2206
|||||
QY 1737 cggcgaagctattctctctcttgatacacactgacttcacacagtatcagacacagaaatac 1796
|||||
Db 2207 TGGGGAAGCAGATCTTCTTCGCAATACATGTACTGTTCCAGTTTCTGATACAGAGAATAC 2266
|||||
QY 1797 ccacgtgtattaaattcttaa 1816
|||||
Db 2257 ACACATCTATTAATCTTTGA 2286
|||||
```

## RESULT 14

```
CTCROMP CTCROMP 2385 bp DNA linear BCT 02-FEB-1993
LOCUS CHlamydia 60kDa cysteine rich outer membrane protein gene.
DEFINITION X53510.1
ACCESSION X53510.1 GI:40681
VERSION outer membrane protein.
KEYWORDS Chlamydia trachomatis.
SOURCE Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
ORGANISM 1 (bases 1 to 2385)
REFERENCE Watson, M.W.
AUTHORS Direct Submission
TITLE Submitted (22-JUN-1990) Watson M.W., Southampton University England,
JOURNAL Department of Microbiology, Faculty of Medicine, South Lab and Path
block, Southampton General Hospital, Tremona road, Southampton, SO9
4XY, UK
REMARK revised by [3]
REFERENCE 2 (bases 43 to 1644)
AUTHORS Watson, M., Lambden, P.R., Ward, M.E. and Clarke, I.N.
TITLE Chlamydia trachomatis 60kDa cysteine rich outer membrane
JOURNAL protein:sequence homology between trachoma and LCV biovars
REFERENCE FEMS Microbiol. Lett. 65, 293-297 (1989)
3 (bases 1 to 2385)
```

## AUTHORS

TITLE Direct Submission  
JOURNAL Submitted (02-FEB-1993) Watson M.W., Southampton University England,  
Department of Microbiology, Faculty of Medicine, South Lab and Path  
block, Southampton General Hospital, Tremona road, Southampton, SO9  
4XY, UK

## COMMENT

FEATURES Data kindly reviewed (06-SEP-1990) by Watson M.

## FEATURES

source Location/Qualifiers  
1..2385  
/organism="Chlamydia trachomatis"  
/db\_xref="taxon:813"  
RBS 199..205  
gene 216..482  
CDS 216..482  
/gene="9kDa Crp"  
/gene="9kDa Crp"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA37587.1"  
/db\_xref="GI:40682"  
/db\_xref="SPTREMBL:O57433"  
/translation="MKKTALLAALCSVLSLSCCRIVDCFCEDPCAPICSPCESKKK  
VDVGCNCNGTVFPACKPCGGDTHQDAKHGPQARGIPVDGRKQ"  
RBS 637..641  
gene 647..2290  
CDS 647..2290  
/gene="60kDa Crp"  
/gene="60kDa Crp"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA37588.1"  
/db\_xref="GI:40683"  
/db\_xref="SWISS-PROT:P18151"  
/translation="MNKLIRRAVTFIAVTSVASLFGVLETSMAESLSTNVISLADT  
KAKNTSHSKKARKNHSHKETPVDRKEVAPVHESKATGPKQKDFVRSOPATPTADCK  
VEITGAVPEYATVSGSPYIEITATGKEDVDVITITQOLPCBAEFVRSOPATPTADCK  
LWKIDRLCOGEKSKINWVKPLKEGCCFTATVCAPEIRSVTHKCGQPAICVQEGP  
ENKLRCPVYIKINIVQGTATKRNWVVPENPFGYAHSSGQGVLFITLGDMPGHRH  
TITVEFCPLKRGATNATVSYCGHKNTASVTVINPCVQVSTKELNDWSTVCKPVE  
YVIVSNPGDLVLRDVEDTLSPGVVLEAQAQISCNKRVVMTKELNDWSTVCKPVE  
LVRAQTPGQFTNNVVKSCDCGCTCAEATYWKGVAAITHMVCVVDTPCVCVGT  
VYRICVTRGSAEDTNVSLMKFSELOPVSFSGPTKGTITGTVVDFSLRIGSKET  
VEFSVTLKAVSAGDARGEAILLSDTLTLPVPSDENTHIY"  
BASE COUNT 708 a 440 c 522 g 715 t

Query Match 42.1%; Score 888; DB 1; Length 2385;  
Best Local Similarity 70.9%; Pred. No. 7.7e-190;  
Matches 1248; Conservative 0; Mismatches 480; Indels 32; Gaps 4;

```
QY 61 aaatgttgagggttaa---gagtttcaaaaacattctaccgatgagcagaagaataaat 116
|||||
Db 566 AAATGTTGAGGGTAAAGTTAGTTAATACCAATTTCTACCCGATGGCAGACAAAAATAA 625
|||||
QY 117 aaacatgcgataggatgccctatgtcccaaaactcatcagcagtagttacggtctctgc 176
|||||
Db 626 TCTATGCGAATAGAGAT-CCTATGAACAAATCATCAGACGACGACGATCTTCGC 684
|||||
QY 177 gctaacagagatgagcagttgcttgcagcgggggtatagagccgctgtacagagtc 236
|||||
Db 685 GGTGACTAGTGTGGCAGGTTTATTGCTAGCGGGTGTAGAGACCTCTATGCCACAGTC 744
|||||
QY 237 tctgattactaagatcgtctcgtcagtcggaagaacaaagccagcactgttctctatgcagc 296
|||||
Db 745 TCTCTCTCAACAGTTATTAGCTTAGCTGACACCA-----AGC 783
|||||
QY 297 gaagaaggttagacttgcctcgttagaaataaacaacccagttgaacaaaaaacgctgtgc 356
|||||
Db 784 GAAAGACACACTTCTCATAAAGCAAAAAACAGAAAAAACCCACAGAACAGACTCC 843
|||||
QY 357 ttttctgataaagaattttatccctctgtaagaggagcagatgtcaacacctgtagagctca 416
|||||
```



```
/protein_id="AAA23154.1"
/db_xref="GI:144558"
/translation="MNLIRRAVTFIAVTSVSLFASGVLETSMAESLSTNVISLADT
KAKDNTSHKSKARKNHSKETLVDREKVPVHESKATGPKQDSFCGRMYTVKVNDRN
VEITQAVPEYATVSGSPYIEITATGKRDVDVIIITQOLPCEAEFVRSOPATPTADGK
LVWKIDRLGCEKESKITVWVKPLKEGECFTAAVCAPERSVTKGQOPALCVKQEGP
ENALRUCPLKRRATNIAVYSCGKHNTASVTIVNEPCQVSLTAGADWVSKPVE
TITVSEFCDLVLVDVVEDTLSPGVLEAAGAOISCNKVVWVVELNPRESLYKV
VVISVNPEDQTNVNVKSCDGCCTSCAEATPYWAGYAATHMVCVDTCDPVCVGEN
VYRICVTGQENADINLSMLKFSKELQVPSFGPTGHTIGNTVFDLPLGSKET
VFESVTLKAVSAGBARGEAILSDTLVFPVSPDENTHIY"
BASE COUNT      712 a 436 c 525 g      717 t
ORIGIN

Query Match      42.1%; Score 888; DB 1; Length 2390;
Best Local Similarity 70.9%; Pred. No. 7.7e-190;
Matches 1248; Conservative 0; Mismatches 480; Indels 32; Gaps 4;

Qy 61 aaatgttgagggtaa----gagttacaaaacattctaccgcatgagcagaagaataat 116
Dy 555 AAATGTTGAGGGTAAAGTTAGTTAATAACAATTCTACCGGATGGCAGACAAAATAA 614
Qy 117 aaacatgcgatagagatccctatgtccaaactcatcagcagtagttagcgtcttgc 176
Dy 615 TCATATGGAATAGGAT--CCTATGAACAAACTCATCAGAGGACGATGACGATCTCGC 673
Qy 177 gctaacagatggcaggttgcctccagcgggggtatagagagcgcgttagcagagtc 236
Dy 674 GGTGACTAGTGTGGCAGGTTATTTCGTAGCGGGGTGTTAGAGACCTCTATGGCAGAGTC 733
Qy 237 tctgattactaagatgcgcagtcagtcggaacaaacagcacctgttctcctatgacgc 296
Dy 734 TCTCTCTACAAAGCTATTAGCTAGCTGACACCAA-----AGC 772
Qy 297 gaagaaggttagactgtccgttagaataaacaacacagttgaacaaaaacgcgtggtgc 356
Dy 773 GAAGACACACCTTCATATAAAGCAAAAAGCAAGAAAACCCACAGCAAGAGACTCT 832
Qy 357 tttttgtataagaattttatccctgtgaagagggagcagatgtcaacctgtagagctca 416
Dy 833 CGTAGACCGTAAAGAGGTTCCTCCGGTTCATGAG-----TCTAAAGCTACAGGACTAA 886
Qy 417 gcaagagctctgtcagcgaagattgtattctgtaaaagtaaacgagatgattgcaacgtaga 476
Dy 887 ACAGGATTCCTGTTTGGCAGAAATGATACAGTCAAGTCAAAAGTTAATGATCATCGTAATCTGA 946
Qy 477 aatttgccagtcgcgttccagaatacgcactactgtaggattccttaccctattgaaatcct 536
Dy 947 AATCACACAGCTGTTCTCTGAATATGCTACGGTAGGATGCCCTATCCCTATTTGAATTC 1006
Qy 537 tgcataaggcaaaaaagattgtgtgtgtgtgattacacacacagctacacctgcgaagc 596
Dy 1007 TGTCTACAGGTAAAGGGATTGTGTGATGTTATCATTTACTACTACAGCAATTACCATTGTGAAGC 1066
Qy 597 tgaattcgtgaagcagtgatccagaaacaaactcctcacagtgatgggaatttagtctggaa 656
Dy 1067 AGAGTTTCGTACGCAGTGTATCCAGCGACAACTCCTACTCTGATGCTAAGCTAGTTTGGAA 1126
Qy 657 aatcgatcgctgggtgcgagagataaatagcataactgtatggttgaataacaccttaa 716
Dy 1127 AATTGACCGCTTAGGACAGGCGAAAGAGTAAATATCTGTATGGGTAAACCTCTTAA 1186
Qy 717 agaaggttgctgtccacagctgctactgtatgtgtgtgctggccagagctcgcgttcttatac 776
Dy 1187 AGAAGGTTGCTGTTTACAGCTGCCAACAGTATGCGCTGTGTCCAGAGATCCGTTCCGGTTAC 1246
Qy 777 taaatgcggttcaaccagcatttgcattaaagcaagaagagacactgactgctgtgcctaaag 836
Dy 1247 AAAATGTGGACAAACCTGCTATCTGTGTTAAACAAGAGGCCAGAGAAATCCTGTTTGGC 1306
Qy 837 atgcccctgtatgctacaaaaatcgaagtagtgaacacagggatctgctattgcccgcgaact 896
```

Search completed: May 25, 2002, 19:55:49  
Job time: 11616 sec



XX WPI: 2000-618918/59.  
DR P-PSDB: AAB18820.  
XX  
XX New polynucleotides encoding a 60kda cysteine-rich membrane protein  
PT from Chlamydia, useful as a vaccine for preventing and treating  
PT Chlamydia infection in mammals -  
XX  
XX Claim 2; Fig 1; 77pp; English.  
XX  
XX The present sequence encodes a Chlamydia 60 kda cysteine-rich membrane  
CC protein. The membrane-rich polynucleotide and polypeptide are useful  
CC for preventing or treating Chlamydia infections, such as community  
CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.  
CC They are also useful for diagnosing Chlamydia infection by assaying a  
CC body fluid of a mammal. The polypeptide is useful for vaccine  
CC production.  
XX  
XX Sequence 2111 BP; 634 A; 417 C; 442 G; 618 T; 0 other;

Query Match 100.0%; Score 2111; DB 21; Length 2111;  
Best Local Similarity 100.0%; Pred No 0;  
Matches 2111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgacaggtagtagagagatgaataattctgactactccttaattcagataataaac 60  
Db 1 ttgacaggtagtagagagatgaataattctgactactccttaattcagataataaac 60  
Qy 61 aaatgttgagggttaagatttacaataattctaccgagtgccagagaataataaac 120  
Db 61 aaatgttgagggttaagatttacaataattctaccgagtgccagagaataataaac 120  
Qy 121 atcgatagagatccctatgccaaactcatcagacagtagttacgctccttgcccta 180  
Db 121 atcgatagagatccctatgccaaactcatcagacagtagttacgctccttgcccta 180  
Qy 181 acagatagtcgaggttgccttcccgagggggtatagagccgctgtagcagagttctcg 240  
Db 181 acagatagtcgaggttgccttcccgagggggtatagagccgctgtagcagagttctcg 240  
Qy 241 attactaagatcgtcgtagtcggaataaacagccagcactgttctctatgacagcgaag 300  
Db 241 attactaagatcgtcgtagtcggaataaacagccagcactgttctctatgacagcgaag 300  
Qy 301 aaggttagactgtcgttagaataataaacacagttgaacaaataaacgctgtgctttt 360  
Db 301 aaggttagactgtcgttagaataataaacacagttgaacaaataaacgctgtgctttt 360  
Qy 361 tggataaagaattttatccctgtgaagagggagcagatgcaacctgtagaggtcagcaa 420  
Db 361 tggataaagaattttatccctgtgaagagggagcagatgcaacctgtagaggtcagcaa 420  
Qy 421 gagtctgtcacggaagatgtattctgtbaaagtaaacgatgattgcaacgtagaatt 480  
Db 421 gagtctgtcacggaagatgtattctgtbaaagtaaacgatgattgcaacgtagaatt 480  
Qy 481 tgcagtcctgtccagaaacacgctactgttaggatctcctaccctattgaaactctgtct 540  
Db 481 tgcagtcctgtccagaaacacgctactgttaggatctcctaccctattgaaactctgtct 540  
Qy 541 ataggcaaaaagattgttattgtattgtattgattacacacagctaccctgcgaagctgaa 600  
Db 541 ataggcaaaaagattgttattgtattgtattgattacacacagctaccctgcgaagctgaa 600  
Qy 601 ttcgtaagcagtgatccagaaacactcctcaagtgatgggaaattagttctggaaatac 660  
Db 601 ttcgtaagcagtgatccagaaacactcctcaagtgatgggaaattagttctggaaatac 660  
Qy 661 gatcgctgggtgcaggaagataaaatgcgaataatttactgtatgggttaaacctcttaagaa 720  
Db 661 gatcgctgggtgcaggaagataaaatgcgaataatttactgtatgggttaaacctcttaagaa 720

Qy 721 ggttgctgtctcaacagctgctactgtatgtcttgcacagagctcogttcttatactaaa 780  
Db 721 ggttgctgtctcaacagctgctactgtatgtcttgcacagagctcogttcttatactaaa 780  
Qy 781 tgcggtcaaccagcatttatttaagcaagaagacctgactgtgcttgcctaaagatgc 840  
Db 781 tgcggtcaaccagcatttatttaagcaagaagacctgactgtgcttgcctaaagatgc 840  
Qy 841 cctgtatgtacataaatacgaagtagtgaacacagatctctattgccctaaacgtact 900  
Db 841 cctgtatgtacataaatacgaagtagtgaacacagatctctattgccctaaacgtact 900  
Qy 901 gtatataactgttcccgatggtctattctcatgcatctggtcgaagagttcttctttt 960  
Db 901 gtatataactgttcccgatggtctattctcatgcatctggtcgaagagttcttctttt 960  
Qy 961 aacttaggagacatgagacctgacataaaaggattattacagttgagttctgcccctaa 1020  
Db 961 aacttaggagacatgagacctgacataaaaggattattacagttgagttctgcccctaa 1020  
Qy 1021 agaagaggtcaaatcaactaactgctactgtaacttactgctgacagctgacacaaatctct 1080  
Db 1021 agaagaggtcaaatcaactaactgctactgtaacttactgctgacagctgacacaaatctct 1080  
Qy 1081 gcaaatgtactacagttgttaagtgccttgcgtacaagtaaatatctctggtgctgat 1140  
Db 1081 gcaaatgtactacagttgttaagtgccttgcgtacaagtaaatatctctggtgctgat 1140  
Qy 1141 tggcttactgctatgtaaacctgtgagtaactctctcatcagtatcgaaatccttgagacttg 1200  
Db 1141 tggcttactgctatgtaaacctgtgagtaactctctcatcagtatcgaaatccttgagacttg 1200  
Qy 1201 gttcttcagatgctgataccaagatacactcctctctggtgttacagtaactcagact 1260  
Db 1201 gttcttcagatgctgataccaagatacactcctctctggtgttacagtaactcagact 1260  
Qy 1261 cctcggggagagatcgtgttaataaagtgttggcggtattcaagaataatggtcccagga 1320  
Db 1261 cctcggggagagatcgtgttaataaagtgttggcggtattcaagaataatggtcccagga 1320  
Qy 1321 gaaacctccagtttaaacctgtagtaaacctcaagtctcttggaagattcacaatactaa 1380  
Db 1321 gaaacctccagtttaaacctgtagtaaacctcaagtctcttggaagattcacaatactaa 1380  
Qy 1381 gttgcagtaactagtagtctaacctgcggaacatgtacatcttgcgagaaacaaca 1440  
Db 1381 gttgcagtaactagtagtctaacctgcggaacatgtacatcttgcgagaaacaaca 1440  
Qy 1441 cattgaaagggtcttgcagctaccatgtgcgtatttagacacaaatgactctatctct 1500  
Db 1441 cattgaaagggtcttgcagctaccatgtgcgtatttagacacaaatgactctatctct 1500  
Qy 1501 gtggggaataactgtctatctgtatctgttaactaacctggttctgtcgaagatct 1560  
Db 1501 gtggggaataactgtctatctgtatctgttaactaacctggttctgtcgaagatct 1560  
Qy 1561 aacgtatctttaaactgtgaagtctcacaagaacttcagccaaatagcttcttcagttcca 1620  
Db 1561 aacgtatctttaaactgtgaagtctcacaagaacttcagccaaatagcttcttcagttcca 1620  
Qy 1621 actaaaggaacatttcaggtaataaccgttcttgcagcctttaccctaactcgtttct 1680  
Db 1621 actaaaggaacatttcaggtaataaccgttcttgcagcctttaccctaactcgtttct 1680  
Qy 1681 aaggaaatctgtagagtttctgttaccctgaaaggtattgtctcccgagagatgctcgcgc 1740  
Db 1681 aaggaaatctgtagagtttctgttaccctgaaaggtattgtctcccgagagatgctcgcgc 1740  
Qy 1741 gaagctatcttcttctgtatcacactgacttccaccagtatcacagacagaaataccac 1800  
Db 1741 gaagctatcttcttctgtatcacactgacttccaccagtatcacagacagaaataccac 1800  
Qy 1801 gtgtattaaattcctaagggaattatcctaaagcagcggaatttccgctctgtcttaggat 1860







Sequence 731 BP: 209 A; 128 C; 158 G; 236 T; 0 other;

[illegible]

RESULT 4  
AAS56989  
ID AAS56989 standard; DNA; 731 BP.  
XX  
XX  
AC  
XX AAS56989;  
XX  
DT 16-JAN-2002 (first entry)  
XX  
XX C. trachomatis DNA clone E6-C8-95 encoding partial 60kDa ORF.  
DE  
XX  
XX Chlamydia; ds; sexually transmitted disease; PID; antibacterial;  
KW pelvic inflammatory disease; antigen; trachoma; gynecological;  
KW acute respiratory tract infection; atherosclerosis; male infertility;  
KW coronary heart disease.

XX	Chlamydia trachomatis serovar E.
OS	
XX	WO200181379-A2.
PN	
XX	01-NOV-2001.
PD	
XX	23-APR-2001; 2001WO-US13081.
XX	
XX	21-APR-2000; 2000US-198853P.
PR	
XX	20-JUL-2000; 2000US-219752P.
XX	
XX	(CORI-) CORIXA CORP.
PA	
XX	Bhatia A, Probst P, Stromberg EJ;
PI	
XX	WPI; 2001-616771/71.
DR	
XX	New polynucleotide for treating Chlamydia infections encodes a
PT	polynucleotides containing an immunogenic portion of a Chlamydia
PT	antigen -
PT	
XX	Claim 1; Page 120; 208pp; English.
XX	
PS	The invention relates to isolated polynucleotide encoding at least
XX	a partial Chlamydia protein which is an antigenic fragment, or the
CC	complements, fragments, homologues and variants, and antibodies
CC	raised against the antigenic proteins (or fragments). The nucleic
CC	acids, proteins and antibodies are used to diagnose and treat Chlamydia
CC	infections (e.g. a sexually transmitted disease, pelvic inflammatory
CC	disease (PID), acute respiratory tract infection, trachoma,
CC	atherosclerosis and coronary heart disease) in a patient, and in
CC	the treatment of male infertility. The compounds of the invention are
CC	also useful for detecting the presence of Chlamydia in a patient, and
CC	stimulating and/or expanding T cells specific for a Chlamydia protein.
CC	The present sequence encodes a Chlamydia antigen.
XX	

Query Match	17.2%	Score 363.6	DB 22	Length 731
Best Local Similarity	75.8%	Pred. No. 4e-88		
Matches 450	Conservative 0	Mismatches 144	Indels 0	Gaps 0
QY 1223	aaagatacactccctctgtgtgtacagtaactgaagctcctgttgagagagctgtcgtga 1282			
Db 4	aagacacitctttctccggagtcagagctcttgaaagctgcaggagctcaaatctctgta 63			
QY 1283	ataaagtgtgtggcgattataaagaatgtgccaggagaaacccctccagtttaaaactg 1342			
Db 64	ataaagtgtgtgactgtgaaagaactgaactctgtgagagctctacagatacaagttc 123			
QY 1343	tagtgaagactcaagttcctggaagattcacaaatcaagttgcagtaactcagtagagctta 1402			
Db 124	tagtaagagcaaaaactcctgacaaattcacaaataaattgttgttgaaagagctgctctg 183			
QY 1403	actgcggaacatgtacatcttgcgcagaaacaaacacacattggaaggctctgcagcta 1462			
Db 184	actgtggtactgttactcttgcgagaagcgacaacttactgtgaaaggagttgctgcta 243			
QY 1463	cccatatgtcgtattagacacaaaatctctatctgtgtaggagaaaaatactgtctatc 1522			
Db 244	ctcatactgtcgttagtagatacttgtgacctgttgtgtaggagaaaaatactgtttacc 303			
QY 1523	gtatctgttaactaaaccgtgtgtctgtgaagatactaaacgtatctttaactctgaagt 1582			
Db 304	gtatttgtgtccaaacagaggttctgcagaagatacaaatgttcttcttaagtcttaaat 363			
QY 1583	tctcaaaagaacttcagccaaatagcttcttcaggtccaaactaaagaagacgatttcagta 1642			
Db 364	tctctaaagaactgtcaacctgtatctctctgtgaccactaaagaacgattacagga 423			
QY 1643	ataccgtgtgttttcgacgctttaccataaactcgtttctaaaggaactctgagagtttctg 1702			
Db 424	atacagtagtattcgattcgtttacctgagtaggttcttaagaagaaactgcagagtttctg 483			
QY 1703	ttacctgaaaggtattgtctcccgagagatgctgcggcggaagctattctttcttcgtata 1762			
Db 484	taacattgaaagcgttacagctggagatgctcgtggggaagcgattcttcttcccgata 543			
QY 1763	cactgacttcacagtatcacacagaaaaatacccacgtgtattataaattcttaa 1816			
Db 544	cattgactgtccagttctgatacagagaatacacacatctattaactcttga 597			
RESULT 5				
AAV16194				
ID	AAV16194 standard; DNA; 230 BP.			
AC	AAV16194;			
XX				
DT	28-MAY-1998 (first entry)			
DE	Part of the gene encoding the 60 kDa cystein rich OMP.			
XX				
XX	Cysteine rich outer-membrane protein; OMP; 60 kDa OMP; 76 kDa OMP;			
KW	hybridisation; amplification; assay; detection; ds.			
XX				
OS	Chlamydia pneumoniae.			
XX				
PN	W09746709-A2.			
XX				
PD	11-DEC-1997.			
XX				
PF	04-JUN-1997; 97WO-US09673.			
XX				
PR	06-JUN-1996; 96US-0659473.			
XX				
XX	(ABBO ) ABBOTT LAB.			
XX				





```
RESULT 9
ID AAA50031 standard: DNA; 800 BP.
XX
AC AAA50031;
XX
DT 10-OCT-2000 (first entry)
XX
DE DNA encoding Chlamydia pneumoniae antigen CPN100696 RY-54.
XX
KW CPN100696 RY-55; antigen; infection; diagnosis; therapy; vaccine;
XX sulfur-rich protein; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
CDS 101..691
FT /*tag= a
XX
PN WO200039158-A1.
XX
PD 06-JUL-2000.
XX
PE 23-DEC-1999; 99WO-CA01230.
XX
PR 23-DEC-1998; 98US-0113280.
XX
PR 23-DEC-1998; 98US-0113281.
XX
PR 23-DEC-1998; 98US-0113282.
XX
PR 23-DEC-1998; 98US-0113283.
XX
PR 23-DEC-1998; 98US-0113284.
XX
PR 23-DEC-1998; 98US-0113285.
XX
PR 23-DEC-1998; 98US-0113385.
XX
PR 28-DEC-1998; 98US-0114050.
XX
PR 28-DEC-1998; 98US-0114056.
XX
PR 28-DEC-1998; 98US-0114057.
XX
PR 28-DEC-1998; 98US-0114058.
XX
PR 28-DEC-1998; 98US-0114059.
XX
PR 28-DEC-1998; 98US-0114061.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Murdin AD, Oomen RP, Wang J;
XX
DR WPI; 2000-452369/39.
XX
PT P-PSDB; AAY95544.
XX
PS Novel Chlamydia polynucleotides and polypeptides useful for diagnosis,
XX prevention and treatment of Chlamydia infection in mammals -
XX
XX Claim 2(a); Page 58-59; 215pp; English.
XX
XX The present sequence is that of Chlamydia pneumoniae genomic DNA
XX including an open reading frame that codes for CPN100696 RY-55 (see
XX AAY95544), a putative sulfur-rich protein. It is an example of
XX C. pneumoniae polynucleotide molecules of the invention (see
XX AAA50030-42) that encode antigenic polypeptides (see AAY95543-55) useful
XX in the diagnosis, treatment and prevention of Chlamydia infection.
XX The polynucleotides can be utilised: in the recombinant production
XX of Chlamydia antigens using transformed unicellular host cells; in
XX vaccines or live vaccine vectors; in naked form or formulated with
XX a delivery vehicle for therapy and prophylaxis of Chlamydia
XX infection; and as probes or primers for diagnosis of Chlamydia
XX infection.
XX
XX Sequence 800 BP; 239 A; 150 C; 162 G; 249 T; 0 other;
XX
XX Query Match 5.7%; Score 120.8; DB 21; Length 800;
XX Best Local Similarity 98.4%; Pred. No. 1.8e-22;
XX Matches 122; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1988 ttattttaaaagcccatcttttagtgatgtaattaaatttttaataaagcttttcta 2047
XX
```

```
Db 1 ttattttaaaagcccatcttttttagtgatgtaattaaatttttaataaagcttttcta 60
Qy 2048 gtgtaacctgtcttttttaggaactacactagggagagacgggtatgtcatcaaatctacatcc 2107
Db 61 gtgtaacctgtcttttttaggaactacactagggagagacgggtatgtcatcaaatctacatcc 120
Qy 2108 cgta 2111
Db 121 cgta 124
RESULT 10
ID AAA70152 standard; DNA; 6033 BP.
XX
AC AAA70152;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:285.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PE 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PR {HOFF//} HOFFMAN S.
XX {CARU//} CARUCCI D.
XX {GARD//} GARDNER M.
XX {VENT//} VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX Plasmodium falciparum, useful as antimalarial vaccines and in the
XX diagnosis of P.falciparum infection -
XX
XX Disclosure; Page 493-495; 577pp; English.
XX
XX The present invention describes proteins and their fragments (I) encoded
XX by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
XX vaccines against P. falciparum infection comprising (I) or (II).
XX (I) and (II) are useful for the development of vaccines against
XX P. falciparum infection. (I) and polyclonal antisera or a monoclonal
XX antibody raised to immunogens comprising the sequences of (I), are
XX useful in the detection of infection with P. falciparum. Furthermore,
XX (I) (especially when they are rifins or secreted or membrane proteins)
XX can aid the identification of drugs to treat or prevent P. falciparum
XX infection, or they can be used to identify drug resistance in
XX P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
XX subsequent identification of proteins encoded by it will help to expand
XX our understanding of parasite biology, a process hampered by the
XX complexity of the parasitic lifecycle, and provide new targets for
XX vaccine and drug development. Parasite resistance to drugs and mosquito
XX resistance to insecticides have led to a resurgence of malaria in many
XX parts of the world, and there is a pressing need for vaccines and new
XX drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
XX and protein sequences given in the present invention, but which are not
XX specifically mentioned within the specification.
XX
XX Sequence 6033 BP; 3019 A; 437 C; 707 G; 1870 T; 0 other;
XX
```

These peptides also induced ICM, indicating that infection with Chlamydia may be involved in the development of ICM. Inflammatory cardiomyopathy peptides are used to determine the risk of ICM by incubation with a subject's T cells and measuring the degree of proliferation (an increased degree being indicative of risk) or to raise specific antibodies which can be used therapeutically and for the detection of Chlamydia. Such peptides can also be used with an adjuvant and an excipient in a vaccine for decreasing ICM.

Sequence 48 BP; 11 A; 10 C; 14 G; 13 T; 0 other;

Query Match 2.2%; Score 46; DB 20; Length 48;  
Best Local Similarity 100.0%; Pred. No. 0.0087;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 ggtagagggccgctgtagcagagtcctctgattactaagatcgtcg 256  
|||||  
DB 1 ggtagagggccgctgtagcagagtcctctgattactaagatcgtcg 46

RESULT 12  
AAZ99167  
ID AAZ99167 standard; DNA; 48 BP.  
XX  
AC AAZ99167;  
XX  
DT 21-JUN-2000 (first entry)  
XX  
DE Chlamydia pneumoniae Cys-rich OMP partial coding sequence.  
XX  
KW Cardiant; murine alpha myosin heavy chain; inflammatory myocarditis;  
KW autoimmune inflammatory cardiomyopathy; Chlamydia; antibody; vaccine;  
KW hybridization probe; ss.  
XX  
OS Chlamydia pneumoniae.  
XX  
FN US6034230-A.  
XX  
PD 07-MAR-2000.  
XX  
PF 03-MAY-1999; 99US-0303862.  
XX  
PR 12-AUG-1998; 98US-0133774.  
XX  
PA (AMGE-) AMGEN CANADA INC.  
XX  
PI Neu N, Penninger JM, Bachmaier K, Hessel AJ;  
XX  
DR WPI: 2000-255712/22.  
XX  
PT P-PSDB; AAY83817.  
XX  
PT DNA molecules encoding novel myocardial peptides used for inhibiting  
XX and inducing inflammatory cardiomyopathy in vivo  
XX  
XX Claim 1; Column 19; 17pp; English.

This sequence represents the coding sequence for a fragment of the Chlamydia pneumoniae Cys-rich outer membrane protein (OMP). The sequence was obtained by homology searches of the PIR database using the murine M7A-alpha/beta peptide consensus sequence (Y83813) as the query. The screen isolated the peptides Y83814-Y83819 and their corresponding coding sequences 299164-299169. The invention relates to the isolation of sequences coding for peptide sequences derived from bacteria and viruses which may cause inflammatory cardiomyopathy. The peptide sequences are searched based on the sequence of the M7A peptides derived from the murine alpha myosin heavy chain polypeptide. The other peptides encoded by the DNAs are used, alone or in conjunction with other therapeutics, for inducing or inhibiting inflammatory cardiomyopathy in vivo, where the cardiomyopathy is autoimmune inflammatory cardiomyopathy, and inflammatory cardiomyopathy caused by Chlamydia or other bacterial or viral infections that cause inflammatory cardiomyopathy. The peptides may also be used for increasing inflammatory

Query Match 2.3%; Score 48; DB 21; Length 6033;  
Best Local Similarity 55.4%; Pred. No. 0.022;  
Matches 93; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1944 ttcttaacaaagatattatttaagttcttgcgaatccttaagatttttttaaaagccca 2003  
|||||  
DB 4366 TTTCGCCCTTTCATATATTATTTATTACTATATACAAATGAAGATTTTCATTATTCGAA 4307

QY 2004 tcttttagtgatgaatttaaaatttttaattagcttcttcctagtgtaacctgtcttt 2063  
|||||  
DB 4306 CATTACATAATTTGTTACTTAAATTTTAAATTTTATTTTAAATGAATTAATCTA 4247

QY 2064 taggaactacactaggagaacggtatgtcatcaaatcatcacccgta 2111  
|||||  
DB 4246 AAGATATAAAGCTAGCAAAAGGAGGTATTTTCATATCTTTTTCGTA 4199

RESULT 11  
AAZ28184  
ID AAZ28184 standard; DNA; 48 BP.  
XX  
AC AAZ28184;  
XX  
DT 20-DEC-1999 (first entry)  
XX  
DE Chlamydia pneumoniae outer membrane protein-derived peptide DNA.  
XX  
KW Heart disease; inflammatory; autoimmune; cardiomyopathy; myosin;  
KW Chlamydia; induction; vaccine; ds.  
XX  
OS Synthetic.  
XX  
OS Chlamydia pneumoniae.  
XX  
FH Key  
FT 1.48  
FT CDS  
FT Location/Qualifiers  
FT /\*tag= a  
FT /product= "Chlamydia pneumoniae outer membrane  
FT protein-derived peptide"  
FT /transl\_except= (pos: 46..48, aa:Ala)

US5962636-A.  
05-OCT-1999.  
12-AUG-1998; 98US-0133774.  
12-AUG-1998; 98US-0133774.  
(AMGE-) AMGEN CANADA INC.  
Bachmaier K, Hessel AJ, Penninger JM, Neu N;  
WPI: 1999-589735/50.  
P-PSDB: AAY42728.  
Peptides that induce or suppress inflammatory cardiomyopathy  
Example 1; Column 19; 17pp; English.

This sequence represents DNA encoding Chlamydia pneumoniae outer membrane protein-derived peptide, which induces inflammatory cardiomyopathy (ICM) via an autoimmune response in mice immunised with it. The peptide contains an amino acid sequence motif Maxxs (AAY42722) which appears to be required for the induction of this disease. The motif was originally identified in the sequence of murine alpha myosin heavy chain-derived peptide, M7A-alpha, when it was compared with a peptide derived from a homologous region of the murine beta myosin heavy chain, M7A-beta (AAY42724) which did not cause the disease. Several peptide fragments containing the motif were identified from a database and were found to be fragments of cysteine rich outer membrane proteins from various species of Chlamydia.

CC myocarditis in a mammal. Antibodies against the peptides and the peptides  
 CC themselves are used for measuring the risk of inflammatory cardiomyopathy  
 CC in a mammal. The peptides may also be used in vaccines. Nucleic acids  
 CC encoding the peptides may be used as hybridization probes, e.g. in  
 CC diagnostic assays to test for the presence of Chlamydia DNA.  
 XX  
 SQ Sequence 48 BP; 11 A; 10 C; 14 G; 13 T; 0 other;

Query Match 2.2%; Score 46; DB 21; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 0.0087;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 211 ggtatagaggcgtgtgacgagctctgtattactaagatcgtcg 256  
 |||||  
 Db 1 ggtatagaggcgtgtgacgagctctgtattactaagatcgtcg 46

RESULT 13  
 ABL33396  
 ID ABL33396 standard; DNA; 5641 BP.  
 XX  
 AC ABL33396;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1369.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.

OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-130909/17.  
 XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 XX cytosine methylation -  
 PS Claim 1; SEQ ID NO 1369; 32pp + Sequence Listing; German.  
 XX

CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX

SQ Sequence 5641 BP; 1442 A; 34 C; 930 G; 3235 T; 0 other;

Query Match 2.1%; Score 43.6; DB 24; Length 5641;  
 Best Local Similarity 100.0%; Pred. No. 0.0087;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 58.5%; Pred. No. 0.33;  
 Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
 QY 1906 ttttttggtttataagctcttcaatcccaatcgtagagttcttcaatcaagatatatt 1965  
 |||||  
 Db 3011 ttttttggtttataagctcttcaatcccaatcgtagagttcttcaatcaagatatatt 1965  
 |||||  
 QY 1966 taagttctgaaatcctaagatttttttaaaagccatcttttaggtatgtaataaa 2025  
 |||||  
 Db 3071 ttgttagtttaattattattttttgattgataaaatttttatttagttattgta 3130  
 |||||  
 QY 2026 attttaatt 2035  
 |||||  
 Db 3131 tattttgatt 3140

RESULT 14  
 ABL33952  
 ID ABL33952 standard; DNA; 7819 BP.  
 XX  
 AC ABL33952;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1925.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.

OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-130909/17.  
 XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 XX cytosine methylation -  
 PS Claim 1; SEQ ID NO 1925; 32pp + Sequence Listing; German.  
 XX

CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX

SQ Sequence 7819 BP; 1808 A; 101 C; 1802 G; 4108 T; 0 other;

Query Match 2.0%; Score 43.2; DB 24; Length 7819;  
 Best Local Similarity 54.4%; Pred. No. 0.49;  
 Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;



```

Db      2061  tataattttatatttttttggatatttttttatatttta 2100
          |  ||   |||    ||||  |  |  |  ||||
Search completed: May 25, 2002, 20:48:39
Job time: 13786 sec
```

Search completed: May 25, 2002, 20:48:39  
Job time: 13786 sec

qy	1906	t t t t t t g t t t t a t a a g c t c t t c a a t c c a a t c g l a g a g t t t c t t a a t c a a g a t a t t t	1965
pb	1941	t g t t t t t t t t t t t a t a t t t t a t t t g t t a t t g a g t t t a t t a t t a t t g g t t t t t t	2000
qy	1966	t a a g t t t t g a a t c t a a g a t t t a t t t t a a a g c c o a t c t t t t t a a g t a t g a t t a a a	2025
pb	2001	t i a g t t t t g t a t t t t t t i a g t t g t a t t t t t t i a g t t t t a a a g t t t t t a t t g g t t t t t t	2060
qy	2026	a t t t t t a a t a a g o t t t t c t a g t a a c c t g t t t t t a	2065
pb	2061	t a t a t t t t t a t t t t t g t g a t a t t t t t a t a t t t a	2100

RESULT	15
ABL34606	
ID	ABL34606 standard; DNA; 7819 BP.
XX	
XX	
AC	ABL34606;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human metastasis associated gene SEQ ID NO: 159.
XX	
XX	Metastasis associated gene; cytostatic; gene therapy; cancer;
KW	cytosine methylation; gene; ds.

Homo sapiens.

XX  
PN WO200177376-A2.

18-OCT-2007  
XX  
XX

XX 3001 3001WQ-EP03970

XX  
XX

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1043826.  
PR 01-SEP-2000; 2000DE-1043826.

XX  
PA  
/EPIC-) EPIGENOMICS AG.XX  
 1. Nienbrock C Berlin K:[illegible][illegible]

pt New nucleic acid derived from chemically useful for diagnosis of cancers by analysis of cytosine methylation

PT 32 also for treatment -

PS Claim 1; SEQ ID NO 159; 23pp + Sequence Listing; English.

The present invention provides a number of human metastasis

genes which are modulated by estrogen used in the diagnosis and treatment of cancer. The present s

one of the genes of the *inv* locus.

SQ Sequence 7819 BP; 1808 A; 101 C; 1802 G; 4108 T; 9 Ccner,

28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841. 842. 843. 844. 845. 846. 847. 848. 849. 850. 851. 852. 853. 854. 855. 856. 857. 858. 859. 860. 8

| Query Match           | Pred. No. | 0.49; |
|-----------------------|-----------|-------|
| Best Local Similarity | 54.48;    |       |

Matches 87; Conservative 0; Nationalism 0

[illegible]

Dd 1941 ttgtttttttttttatatttttttatttggtattgatctatccacaggtc

1966 taagtttctgaaatcctaagatttatTTTaaagcccatcttttttaggtatgttaa

nb 2001 ttagtttttqtatttttttagttgtatttttttagttttaagttttttatttggtt

2025 +++++aagcttttcctagtgtaacctgcttcttta 2065

17



```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on:      May 25, 2002, 16:54:33 ; Search time 89.21 Seconds
              (without alignments)
              5812.491 Million cell updates/sec

Title:      US-09-523-647-1
Perfect score: 2111
Sequence:    1 ttgatcaggtagttaggaga.....catcaaatctacatccgta 2111

Scoring table:
Gapop 10.0 , Gapext 1.0

```

Searched: 383533 seqs, 122816752 residues

## SUMMARIES

| Result No. | Score | Query |       | Length | DB               | ID                 | Description |
|------------|-------|-------|-------|--------|------------------|--------------------|-------------|
|            |       | Match | %     |        |                  |                    |             |
| 1          | 228.4 | 10.8  | 230   | 4      | US-08-659-473-1  | Sequence 19, Appli |             |
| 2          | 173.4 | 8.2   | 327   | 1      | US-08-477-270-19 | Sequence 20, Appl  |             |
| 3          | 46    | 2.2   | 48    | 2      | US-09-133-774-20 | Sequence 20, Appl  |             |
| 4          | 46    | 2.2   | 48    | 3      | US-09-303-862-20 | Sequence 14, Appl  |             |
| c 5        | 45.6  | 2.2   | 7218  | 1      | US-08-232-463-14 | Sequence 31, Appl  |             |
| c 6        | 39.6  | 1.9   | 1508  | 3      | US-08-714-918-31 | Sequence 31, Appl  |             |
| c 7        | 39.6  | 1.9   | 1508  | 4      | US-09-265-315-31 | Sequence 31, Appl  |             |
| c 8        | 39.6  | 1.9   | 1508  | 4      | US-09-265-315-31 | Sequence 31, Appl  |             |
| c 9        | 38.6  | 1.9   | 1508  | 4      | US-09-266-417-31 | Sequence 31, Appl  |             |
| c 10       | 38.6  | 1.8   | 3360  | 3      | US-08-938-546-5  | Sequence 5, Appli  |             |
| c 11       | 38.6  | 1.8   | 3360  | 3      | US-09-340-812-5  | Sequence 5, Appli  |             |
| c 12       | 38.2  | 1.8   | 72328 | 3      | US-09-009-913-1  | Sequence 1, Appli  |             |
| c 13       | 37.8  | 1.8   | 2739  | 3      | US-08-678-613-1  | Sequence 1, Appli  |             |
| c 14       | 37.8  | 1.8   | 2902  | 3      | US-08-714-918-95 | Sequence 95, Appl  |             |
| c 15       | 37.8  | 1.8   | 2902  | 4      | US-09-265-315-95 | Sequence 95, Appl  |             |
| c 16       | 37.8  | 1.8   | 2902  | 4      | US-09-265-315-95 | Sequence 95, Appl  |             |
| c 17       | 37.8  | 1.8   | 2902  | 4      | US-09-266-417-95 | Sequence 95, Appl  |             |
| c 18       | 37.6  | 1.8   | 824   | 1      | US-08-158-353-1  | Sequence 1, Appli  |             |
| c 19       | 36.8  | 1.7   | 1189  | 1      | US-07-781-034-4  | Sequence 4, Appli  |             |
| c 20       | 36.8  | 1.7   | 1189  | 5      | PCT-US92-08328-4 | Sequence 4, Appli  |             |
| c 21       | 36.8  | 1.7   | 1365  | 4      | US-09-308-003-3  | Sequence 3, Appli  |             |
| c 22       | 36    | 1.7   | 972   | 3      | US-09-286-690-1  | Sequence 3, Appli  |             |
| c 23       | 35.6  | 1.7   | 1151  | 1      | US-08-236-754-3  | Sequence 3, Appli  |             |
| c 24       | 35.6  | 1.7   | 2476  | 4      | US-08-844-274-11 | Sequence 11, Appli |             |
| c 25       | 35.6  | 1.7   | 3055  | 1      | US-08-236-754-1  | Sequence 1, Appli  |             |
| c 26       | 35.6  | 1.7   | 5194  | 4      | US-08-844-274-16 | Sequence 16, Appl  |             |
| c 27       | 35.6  | 1.7   | 5194  | 4      | US-08-844-274-17 | Sequence 17, Appl  |             |

```
QY 533 tccttgctataggcaaaaagattgctgttgattgattacacacagctacottgq 592
Db 61 TCCTTGCTATAGGCAAAAAGATTGTTGTTGCTGATTACACACACCTACCTTGGC 120
QY 593 aagctgaattcgaagcagtgatcagaaacaaactcctcacaagtgatgggaaattagtct 652
Db 121 AAGCTGAATTCGTAAAGCAGTATCCAGAAACAACTCCTTACAAAGTGATGGGAAATTAGTCT 180
QY 553 ggaataatgatcgctcgctgggtcagagagataaaatgcaaaattactgtatgg 702
Db 181 GGAAAATCGATCGCTGGTGCAGGAGATAAATCCAAAATTACTGTATGG 230

RESULT 2
US-08-477-270-19
; Sequence 19, Application US/08477270
; Patent No. 5629158
; GENERAL INFORMATION:
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL
; TITLE OF INVENTION: CONDITIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,270
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,010
; FILING DATE:
; APPLICATION NUMBER: US 07/781,157
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/153 DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..327
; FEATURE:
; NAME/KEY: misc_difference
; LOCATION: replace(60,"")
; OTHER INFORMATION: /note="The C at position 60 can
; OTHER INFORMATION: also be a T. The amino acid at position 20 will
; OTHER INFORMATION: be Asn in either case."
; FEATURE:
; NAME/KEY: misc_difference
; LOCATION: replace(178,"")
; OTHER INFORMATION: /note="The A at position 178 can
; OTHER INFORMATION: also be a G. If it is a G, then the amino acid at
```

```
; OTHER INFORMATION: position 60 Glu instead of Lys."
US-08-477-270-19
```

```
Query Match 8.2%; Score 173.4; DB 1; Length 327;
Best Local Similarity 70.6%; Pred. No. 2.6e-41;
Matches 231; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1045 gctactgaacttactcggtggacacaaatgttctgcacaaatgtaactacagttgttaat 1104
Db 1 GCATGTTCTTCTACTGTGGAGGACATAAAATAATACAGCAAGGTAACAACTGTGATCAAC 60

QY 1105 gagccttggtacaaataatctctggtgctgtcttctacgtatgttaaacctgtg 1164
Db 61 GAGCCTTTCGCTACAAAGTAAGTATTGCGAGGACAGATTTGGTCTTATGTTTGAAGCCTGTA 120

QY 1165 ggtactctatctcagtatcgaatcctggagacttggttcttctcatgatgctgataccaa 1224
Db 121 GAATATGTGATCTCCGTTTCCAATCTGTGGAGATCTTGTGTCCGAGATGTCGTCGTTAAA 180

QY 1225 gatactcctctctggtgttacagtaactcgaagctcctcgtgagagatctgctgtaat 1284
Db 181 GACACTCTTTCTCCCGGAGTCACAGTTCTTTGAAGCTGCAGAGCTCAAAATTTCTTGTGAAT 240

QY 1285 aaagttgttgcggtattaaagaaatgtgccaggagaaacccctccagtttaaacctgta 1344
Db 241 AAAGTAGTTGGACTGTGAAAGAACTGAATCCTGGAGAGTCTCTACAGTATAAAGTTCTA 300

QY 1345 gtgaagagctcaagttcctcgtgaagattc 1371
Db 301 GTAAGAGACACAACTCCTCGACAAATTC 327
```

```
RESULT 3
US-09-133-774-20
; Sequence 20, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Hear
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-133-774-20
```

```
Query Match 2.2%; Score 46; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 ggtatagagccgctgtgacagagctctctgattactaagatcgtcg 256
Db 1 ggtatagagccgctgtgacagagctctctgattactaagatcgtcg 46
```

```
RESULT 4
US-09-303-862-20
; Sequence 20, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
```



ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 222/005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1508 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-714-918-31

Query Match 1.9%; Score 39.6; DB 3; Length 1508;  
Best Local Similarity 49.5%; Pred. No. 0.08;  
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 503 ctactgtaggatctcctaccattgaaatccttgcctataggaacaaagattgtgttg 562  
Db 1268 CAATTGAAGATTTTCATTTAATAAGAAATTTTCGAAAAAGAAATTAGAGGCTTTA 1209  
QY 563 atgtgtgattacacacagctaccttgcgaagctgaattctgaagcagtgatccagaaa 622  
Db 1208 ATATCACAATTCCTATAAAGAACGTATCATACCGTATTATGATTATGATGAACAAG 1149  
QY 623 caactcctacagtgatgggaaattagctctggaaatcgatcgctgggtgcagagata 682  
Db 1148 CGATTAATGCAGGTGCAGTTAACACTGTTTGTGATAAAGATGGCAAGTGGATAGGTATA 1089  
QY 683 aatgcaaaattactgtatgggtataaa 708  
Db 1088 ATACAGATGGTATGGTTATGTTAA 1063

RESULT 7  
US-09-265-315-31/C  
Sequence 31, Application US/09265315  
Patent No. 6187541  
GENERAL INFORMATION:  
APPLICANT: Benton, Bret  
APPLICANT: Lee, Ving J.  
APPLICANT: Malouin, Francois  
APPLICANT: Martin, Patrick K.  
APPLICANT: Schmid, Molly B.  
APPLICANT: Sun, Dongxu  
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
FILING DATE: March 9, 1999  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996  
APPLICATION NUMBER: 60/009,102  
FILING DATE: December 22, 1995  
APPLICATION NUMBER: 60/003,798  
FILING DATE: September 15, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 240/247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1508 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-265-315-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;  
Best Local Similarity 49.5%; Pred. No. 0.08;  
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 503 ctactgtaggatctcctaccattgaaatccttgcctataggaacaaagattgtgttg 562  
Db 1268 CAATTGAAGATTTTCATTTAATAAGAAATTTTCGAAAAAGAAATTAGAGGCTTTA 1209  
QY 563 atgtgtgattacacacagctaccttgcgaagctgaattctgaagcagtgatccagaaa 622  
Db 1208 ATATCACAATTCCTATAAAGAACGTATCATACCGTATTATGATTATGATGAACAAG 1149  
QY 623 caactcctacagtgatgggaaattagctctggaaatcgatcgctgggtgcagagata 682  
Db 1148 CGATTAATGCAGGTGCAGTTAACACTGTTTGTGATAAAGATGGCAAGTGGATAGGTATA 1089  
QY 683 aatgcaaaattactgtatgggtataaa 708  
Db 1088 ATACAGATGGTATGGTTATGTTAA 1063

RESULT 8  
US-09-265-315-31/C  
Sequence 31, Application US/09265315  
Patent No. 6187541  
GENERAL INFORMATION:  
APPLICANT: Benton, Bret  
APPLICANT: Lee, Ving J.  
APPLICANT: Malouin, Francois  
APPLICANT: Martin, Patrick K.  
APPLICANT: Schmid, Molly B.  
APPLICANT: Sun, Dongxu  
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:

Tue May 28 08:55:08 2002

APPLICATION NUMBER: US/09/265,315  
FILING DATE: March 9, 1999  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/714,918  
FILING DATE: September 13, 1996  
APPLICATION NUMBER: 60/009,102  
FILING DATE: December 22, 1995  
APPLICATION NUMBER: 60/003,798  
FILING DATE: September 15, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 240/247  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1508 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-265-315-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;  
Best Local Similarity 49.5%; Pred. No. 0.08;  
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctcttaccctattgaaatccttgcctataggcgaagaaagattgtgtg 562  
DB 1268 CAATTGAAGATTTTCATTTAAATAAGAAATTTTCGAAAAAAGAAATTTAGAAAGCTTTA 1209

QY 563 atgttgatgattacacacagctaccttgcgaagctgaattctgaagcagctgacccagaaa 622  
DB 1208 ATATCACAAATTCCTCAATAAGAACGTATACCGTATTAGATTATGTTGATGAACAAG 1149

QY 623 caactcctacaagtgatggaaattagctctggaataatcgcctcgtggtgagagagata 682  
DB 1148 CGATTAATGCAGTGCAGTTAACACTGTTTGTATAAAGATGCAAGTGGATGGGTATATA 1089

QY 683 atgcgaataattactgtatgggtaaaa 708  
DB 1088 ATACAGATGTTATGTTTATGTTTAAA 1063

RESULT 9  
US-09-266-417-31/c  
Sequence 31, Application US/09266417  
Patent No. 6228588  
GENERAL INFORMATION:  
APPLICANT: Benton, Bret  
APPLICANT: Lee, Ving J.  
APPLICANT: Malouin, Francois  
APPLICANT: Martin, Patrick K.  
APPLICANT: Schmid, Molly B.  
APPLICANT: Sun, Dongxu  
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
TITLE OF INVENTION: TARGET GENES  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/266,417  
FILING DATE: March 9, 1999  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/714,918  
FILING DATE: September 13, 1996  
APPLICATION NUMBER: 60/009,102  
FILING DATE: December 22, 1995  
APPLICATION NUMBER: 60/003,798  
FILING DATE: September 15, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 240/248  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1508 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-266-417-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;  
Best Local Similarity 49.5%; Pred. No. 0.08;  
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctcttaccctattgaaatccttgcctataggcgaagaaagattgtgtg 562  
DB 1268 CAATTGAAGATTTTCATTTAAATAAGAAATTTTCGAAAAAAGAAATTTAGAAAGCTTTA 1209

QY 563 atgttgatgattacacacagctaccttgcgaagctgaattctgaagcagctgacccagaaa 622  
DB 1208 ATATCACAAATTCCTCAATAAGAACGTATACCGTATTAGATTATGTTGATGAACAAG 1149

QY 623 caactcctacaagtgatggaaattagctctggaataatcgcctcgtggtgagagagata 682  
DB 1148 CGATTAATGCAGTGCAGTTAACACTGTTTGTATAAAGATGCAAGTGGATGGGTATATA 1089

QY 683 atgcgaataattactgtatgggtaaaa 708  
DB 1088 ATACAGATGTTATGTTTATGTTTAAA 1063

RESULT 10  
US-08-938-546-5/c  
Sequence 5, Application US/08938546  
Patent No. 6004556  
GENERAL INFORMATION:  
APPLICANT: Fosberry, Andrew P.  
APPLICANT: Lawlor, Elizabeth J.  
APPLICANT: Nicholas, Richard O.  
TITLE OF INVENTION: No. 6004556el rsbu-1  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

```

RESULT 11
US-09-340-812-5/c
; Sequence 5, Application US/09340812
; Patent No. 611078
; GENERAL INFORMATION:
; APPLICANT: Fosberry, Andrew P.
; APPLICANT: Lawlor, Elizabeth J.
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: NO. 611078el rsbu-1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq for Windows Version 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/340,812
; FILING DATE:
; CLASSIFICATION:

```

```

RESULT 12
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: AXYS Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:

```



TELEPHONE: 650-327-3231  
TELEFAX: 650-327-3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-009-913-1

Query Match 1.8%; Score 38.2; DB 3; Length 72928;  
Best Local Similarity 50.8%; Pred. No. 1.6;  
Matches 91; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
QY 1882 ttgtacctacgtactaaagcggtttttgtttataagctcttcaatccaatcgtag 1941  
DB 64608 TTGTTTCTCGACAGCTTAGTAGCTTGGCCCTTTGAAACCCCGTTCAATTTAAGTAAC 64549  
QY 1942 agttcttaataaagattatttaagttctgaaatcctaaagattatttttaaagcc 2001  
DB 64548 ACGTTTTTAATACAGCAAGATTGTTGCTCTTTAGACAAATCTATTATTTAAATCA 64489  
QY 2002 catcttttaggtatgtaataaaattttaataagctttctcctagtgtaacctgctt 2060  
DB 64488 TATCTTAAATATAGATATGAAAAATGTCAAAATAAAGTTAAATTATCTATGCCCCCTCTT 64430

RESULT 13  
US-08-678-613-1/c  
; Sequence 1, Application US/08678613  
; Patent No. 6136587  
; GENERAL INFORMATION:  
; APPLICANT: Tomasz, Alexander  
; APPLICANT: De Lencastre, Herminia  
; TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF  
; TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/678,613  
; FILING DATE: 10-JUL-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,319  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-140  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2739 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus aureus  
; STRAIN: RUSA 168  
; US-08-678-613-1

Query Match 1.8%; Score 37.8; DB 3; Length 2739;  
Best Local Similarity 52.9%; Pred. No. 0.37;  
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 1873 agtaccgcttttagctacgtactaaagcggtttttgtttataaagctcttcaatc 1932  
DB 2093 AATACGCTTCATTTGTTTTTACTAAATGATATTTCTCTATGATTGAATTAAATCAAT 2034  
QY 1933 caatcgtagaggtttcttaacaaagattatttaagtttctgaaatcctaaagatttatt 1992  
DB 2033 TGATTGTCCATTTACTTTAACCCACACTTTTCTTTAAATTTACTCATTTATTTTAAACATAT 1974  
QY 1993 ttaaagcccatcttttaggtatgtaataaaa 2025  
DB 1973 TCTAAATACTTCTATTAAAGGTATGATACTTAA 1941

RESULT 14  
US-08-714-918-95/c  
; Sequence 95, Application US/08714918  
; Patent No. 6037123  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: Los Angeles  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,918  
; FILING DATE: September 13, 1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 222/005  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2902 base pairs



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: May 25, 2002, 16:51:56 ; Search time 2937.23 Seconds  
(without alignments)  
9700.320 Million cell updates/sec

Title: US-09-523-647-1  
Perfect score: 2111  
Sequence: 1 ttgtacagtagtaggaga.....catcaaatctacatccogta 2111

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1          | 57.2  | 2.7           | 1101   | 12 | CNS0039G    |
| 2          | 51.4  | 2.4           | 1101   | 12 | CNS0039G    |
| 3          | 51.2  | 2.4           | 1173   | 12 | CNS06LVV    |
| 4          | 50.8  | 2.4           | 928    | 12 | CNS00DKY    |
| 5          | 50.2  | 2.4           | 1201   | 12 | CNS016E1    |
| 6          | 48.6  | 2.3           | 1101   | 12 | CNS016DT    |
| 7          | 47    | 2.2           | 586    | 12 | AZ323116    |
| 8          | 46.6  | 2.2           | 698    | 12 | CNS010V2    |
| 9          | 45.8  | 2.2           | 976    | 12 | CNS04F5M    |
| 10         | 45.2  | 2.1           | 610    | 9  | AA675942    |
| 11         | 45.2  | 2.1           | 686    | 9  | AL514705    |
| 12         | 45.2  | 2.1           | 822    | 12 | CNS009CW    |
| 13         | 45.2  | 2.1           | 987    | 12 | CNS014PQ    |
| 14         | 45    | 2.1           | 1101   | 12 | CNS001FB    |
| 15         | 44.4  | 2.1           | 1205   | 12 | CNS0165A    |
| 16         | 44    | 2.1           | 801    | 12 | BH499269    |
| 17         | 43.8  | 2.1           | 843    | 12 | CNS00C51    |

|    |      |     |      |    |          |
|----|------|-----|------|----|----------|
| 18 | 43.8 | 2.1 | 987  | 12 | CNS014PQ |
| 19 | 43.6 | 2.1 | 975  | 12 | CNS0102X |
| 20 | 43.4 | 2.1 | 1225 | 12 | CNS016ID |
| 21 | 43.2 | 2.0 | 945  | 12 | CNS013W4 |
| 22 | 43   | 2.0 | 380  | 9  | AW240912 |
| 23 | 43   | 2.0 | 843  | 12 | CNS00C51 |
| 24 | 43   | 2.0 | 857  | 12 | BH552085 |
| 25 | 43   | 2.0 | 1001 | 12 | CNS0155H |
| 26 | 42.8 | 2.0 | 658  | 9  | BA438151 |
| 27 | 42.8 | 2.0 | 924  | 12 | CNS07A5L |
| 28 | 42.8 | 2.0 | 1135 | 12 | CNS033GQ |
| 29 | 42.6 | 2.0 | 1101 | 12 | CNS00DT7 |
| 30 | 42.6 | 2.0 | 1101 | 12 | CNS017KT |
| 31 | 42.6 | 2.0 | 1102 | 12 | CNS014P4 |
| 32 | 42.4 | 2.0 | 802  | 12 | CNS0383B |
| 33 | 42.4 | 2.0 | 1101 | 12 | CNS002AL |
| 34 | 42.2 | 2.0 | 623  | 12 | AQ288608 |
| 35 | 42.2 | 2.0 | 945  | 12 | CNS04D0K |
| 36 | 42.2 | 2.0 | 1092 | 12 | CNS020K7 |
| 37 | 42.2 | 2.0 | 1168 | 12 | CNS06HDF |
| 38 | 42   | 2.0 | 531  | 10 | BG602527 |
| 39 | 42   | 2.0 | 1101 | 12 | CNS00LT2 |
| 40 | 41.8 | 2.0 | 298  | 12 | B22209   |
| 41 | 41.8 | 2.0 | 367  | 12 | CNS00046 |
| 42 | 41.8 | 2.0 | 472  | 10 | BG226909 |
| 43 | 41.8 | 2.0 | 578  | 12 | AQ010923 |
| 44 | 41.8 | 2.0 | 961  | 12 | CNS008H1 |
| 45 | 41.8 | 2.0 | 1061 | 12 | CNS015LM |

ALIGNMENTS

RESULT 1  
CNS0039G/C  
LOCUS  
DEFINITION  
CNS0039G 1101 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BAC08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921 GI:4941778  
VERSION  
AL063921.1  
KEYWORDS  
GSS.  
SOURCE  
fruit fly.  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 1101)  
Genoscope.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
JOURNAL  
- Web : www.genoscope.cns.fr)

COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's Pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES

1..1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCI-98"

```

/clone="BACR08K10"
/notes="end : TET3"
BASE COUNT      201 a      64 c      131 g      202 t      503 others
ORIGIN

Query Match      2.7%; Score 57.2; DB 12; Length 1101;
Best Local Similarity 17.2%; Pred. No. 0.0018;
Matches 123; Conservative 275; Mismatches 318; Indels 0; Gaps 0;

QY 1348 aagaactcctcgaagattcacaaatacaagttgcagtaactagtgactcaactgc 1407
Db 1081 DWTGTTTWWKDRDRRWAGDADRWADGAGTWTATWMMWMAATWDTWDKWWWA 1022

QY 1408 ggaacatgacattcgtcgagaaacacacacattggaagttcgtcagctaccatc 1467
Db 1021 TAAKTDTAWTWTAWRADRGAGKRDRAATDADAGRDKGKRKDKDRDXGD 962

QY 1468 atgctgattacacaaatgctctctctgtgtgaggagaaatactgtctcgtatc 1527
Db 961 DKKGGKKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 902

QY 1528 tgtgttaactaacctggtctgctgagatacactaacgtatctttaaactcgtc 1587
Db 901 DDTGTDKDDDKDKWDDKAKCTGDAWAAATDWWMGWADADWTTWDAADADWAD 842

QY 1588 aaagaacttcagcaaatgctctcagtcacaaactaaaggaacgattcagtaatac 1647
Db 841 DRWDAAWAKWDDAAWAGARTADRRDWRGDRAGKRGKARRKRRKRRKRRKRR 782

QY 1648 gttgtttcagcgtttacataactcgttcttaaggaactcgtagagtttctgtac 1707
Db 781 WTTTWTTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRD 722

QY 1708 tgaagattgctcccgagagctcgcgcggaagctattcttctctcgtcacactg 1767
Db 721 DTRWDADADDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRD 662

QY 1768 actcaccagttatcagacagaaataccacgcggtattaaattcgaagaaattctc 1827
Db 661 TDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRD 602

QY 1828 aaagcagagcagattccgctcgtctgttagagctctcgaagaaagcagctcttag 1887
Db 601 TDARKADRWAKARAWRADRRARARADRRWTKGTTTATTTTAAARAAWAAWAT 542

QY 1888 cctcagtaactaagcgtttttttgtttataagctctcgaactcgaactagagttc 1947
Db 541 TTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 482

QY 1948 ttaatcaagatatatttaagttctgaaatcctcgaagatttttttaaaagccactct 2007
Db 481 TTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 422

QY 2008 tttaggtatgaataaaatttttaagcttttctcgttagtaacctgtcttct 2063
Db 421 WTTTWTAAAAAAMWMTWTATATKCCCCCYCCCCCYCCCCCYCCCTCCYCCCT 366

```

## RESULT 2

CNS0039G

LOCUS

DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL063921

VERSION AL063921.1 GI:4941778

KEYWORDS GSS;

SOURCE fruit fly.

ORGANISM Drosophila

taxonomy: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1. (bases 1 to 1101)

Genoscope.

Direct Submission

AUTHORS

TITLE

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library.html](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library.html) was prepared by Kazuhiro Oseogawa and Aaron Mammeter in Pictet de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone\_lib="RPc1-98"

/clone="BACR08K10"

/note="end : TET3"

BASE COUNT 201 a 64 c 131 g 202 t 503 others

ORIGIN

Query Match

Best Local Similarity 2.4%; Score 51.4; DB 12; Length 1101;

Matches 84; Conservative 263; Mismatches 244; Indels 0; Gaps 0;

```

QY 1475 tattagacacaaatgactctatctgtgtgaggagaaatactgtctatcgtgtaa 1534
Db 510 WTTTAAWAAAAAANAANAANAATAATATWTTTWTWTAAWATAAAMCAA 569

QY 1535 ctacccgtggtctgctgaagatacactacgtatctttaaactcgtcgtcctcaagaac 1594
Db 570 WYHTTWTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTY 629

QY 1595 ttcagccaaatagctcttcaggtcccaactaaaggaacattcaggtacacgttctt 1654
Db 630 WHWTYTAAYYYTYCYYYHYMHHAHAHAHAHAHAHAHAHAHAHAHAHAHA 689

QY 1655 tcgacgtttacacaaactcgttcttaaggaactcgttagagttctcgtcctcgaag 1714
Db 690 MMCCTTCTCHCYYYHYHTAHHTHTHHWYAHYMYMYMYMYMYMYMYMYMY 749

QY 1715 gtattgctccgagagctcgcgcggaagctattcttctcgtacactgactctc 1774
Db 750 HWYHTTWYAWAHAMWMMHHHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 809

QY 1775 cagatcacagacagaaataccacagctgtattaaattcgaagaaattcctcaagcag 1834
Db 810 MCTYHCWHYHTAYTCWTWTHHMMWTWTHWYHTTWHHTTTHAAWHTTTCMCHHAT 869

QY 1835 agcgatattccgctcgtcttagagatagcttccaaagaagctcgtcttagctacc 1894
Db 870 WTWATCWCACMTMHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 929

QY 1895 tactaaagcgtttttttgtttttaaagctctcctcaatcgaatcgtagagttctt 1954
Db 930 MHHHHHHHHHHHMAATWMTTMTTMMHMMHCHHCHHCHHCHHCHHCHHCHH 989

QY 1955 aagattatttaagttctgaaatccttaagatttattttaaaagccactcttctt 2014
Db 990 YHYMCTCHYCTWHTYTYWYAWTAHAMTTATWMMHMMHMAHAMATWMMHMA 1049

QY 2015 atgtaataaaatttttaagcttttctcgtagtgtaacctgctcttctt 2065

```



```

RESULT 5
CNS016E1/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN15C13 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106627
VERSION AL106627.1 GI:5622848
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES             Location/Qualifiers
     source            1..1201
                        /organism="Drosophila melanogaster"
                        /plasmid="pBelOBAC11"
                        /db_xref="taxon:7227"
                        /clone_lib="DrosBAC"
                        /clone="BACN15C13"
                        /note="end : SP6"
BASE COUNT 448 a 136 c 161 g 336 t 120 others
ORIGIN

Query Match      2.4%; Score 50.2; DB 12; Length 1201;
Best Local Similarity 41.2%; Pred. No. 0.1;
Matches 96; Conservative 30; Mismatches 107; Indels 0; Gaps 0;

QY 1867 caaagaagtcgcttagtaacctacgtactaaagcggtttttgttttataagctct 1926
Db 1189 MMAMAMMMMHMTCTTTTATATATAAAAAAAWATTWTWTWTWTWTWTWTAAWCT 1130

QY 1927 tcaatccaatcgtagagttcttaatacaagatatattaaagttctgaaatcctaaga 1986
Db 1129 TTTTATTATATATATATATATATATATATATATATATATATATATATATATAT 1070

QY 1987 ttattttaaagcccatcttttttaggtatgtaataaaatttttaataagctttct 2046
Db 1069 TTTTWTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1010

QY 2047 agtgaacctgctcttttaggaactacactaggagaacggtatgctcaaat 2099
Db 1009 TTTTATTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 957

RESULT 6
CNS016D7/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN15B04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106619
VERSION AL106619.1 GI:5622820
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES             Location/Qualifiers
     source            1..1101
                        /organism="Drosophila melanogaster"
                        /plasmid="pBelOBAC11"
                        /db_xref="taxon:7227"
                        /clone_lib="DrosBAC"
                        /clone="BACN15B04"
                        /note="end : T7"
BASE COUNT 366 a 129 c 148 g 305 t 153 others
ORIGIN

Query Match      2.3%; Score 48.6; DB 12; Length 1101;
Best Local Similarity 39.1%; Pred. No. 0.25;
Matches 70; Conservative 38; Mismatches 71; Indels 0; Gaps 0;

QY 1906 ttctttgttttataagctcttcaatccaatcgtagagttcttcaatacaagattatt 1965
Db 994 HHHTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 935

QY 1966 taagttctgaaatcctaagattattttaaagcccatcttttaggtatgataataaa 2025
Db 934 AWAATTTTTCAMCTTMMWAAWTTTMMWAAWTTTMMWAAWTTTMMWAAWTTTMMWAA 875

QY 2026 atttcaataagctttcttagtgaacctgcttctttaggaactacactaggagaac 2084
Db 874 WTATATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 816

RESULT 7
AZ323116/c
LOCUS
DEFINITION 1M0044506F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
clone UUGC1M0044G06 F, DNA sequence.
ACCESSION AZ323116
VERSION AZ323116.1 GI:10377520
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 586)
AUTHORS Dunn,D., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

```

Insert Length: 10000    Std Error: 0.00  
Plate: 0044    row: G    column: 06  
Seq primer: CGTTGTAACACACGGCCAGT  
Class: plasmid ends  
High quality sequence stop: 586.

| FEATURES              | SOURCE    |
|-----------------------|-----------|
| 1. <b>Feature 1</b>   | Source 1  |
| 2. <b>Feature 2</b>   | Source 2  |
| 3. <b>Feature 3</b>   | Source 3  |
| 4. <b>Feature 4</b>   | Source 4  |
| 5. <b>Feature 5</b>   | Source 5  |
| 6. <b>Feature 6</b>   | Source 6  |
| 7. <b>Feature 7</b>   | Source 7  |
| 8. <b>Feature 8</b>   | Source 8  |
| 9. <b>Feature 9</b>   | Source 9  |
| 10. <b>Feature 10</b> | Source 10 |

I. 586  
 /organism="Mus musculus"  
 /strain="C57BL/6J  
 /db\_xref="taxon:10090"  
 /clone="UGC1M0044606"  
 /clone\_lib="Mouse 10kb plasmid UGC1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="vector: pWD42nv; Purified genomic DNA from F.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (<http://www.jax.org/resources/documents/dnares/>). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gll4732114[gblAF129072.1], a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

|     | BASE COUNT | ORIGIN |
|-----|------------|--------|
| 1   | 0          | 0      |
| 2   | 0          | 0      |
| 3   | 0          | 0      |
| 4   | 0          | 0      |
| 5   | 0          | 0      |
| 6   | 0          | 0      |
| 7   | 0          | 0      |
| 8   | 0          | 0      |
| 9   | 0          | 0      |
| 10  | 0          | 0      |
| 11  | 0          | 0      |
| 12  | 0          | 0      |
| 13  | 0          | 0      |
| 14  | 0          | 0      |
| 15  | 0          | 0      |
| 16  | 0          | 0      |
| 17  | 0          | 0      |
| 18  | 0          | 0      |
| 19  | 0          | 0      |
| 20  | 0          | 0      |
| 21  | 0          | 0      |
| 22  | 0          | 0      |
| 23  | 0          | 0      |
| 24  | 0          | 0      |
| 25  | 0          | 0      |
| 26  | 0          | 0      |
| 27  | 0          | 0      |
| 28  | 0          | 0      |
| 29  | 0          | 0      |
| 30  | 0          | 0      |
| 31  | 0          | 0      |
| 32  | 0          | 0      |
| 33  | 0          | 0      |
| 34  | 0          | 0      |
| 35  | 0          | 0      |
| 36  | 0          | 0      |
| 37  | 0          | 0      |
| 38  | 0          | 0      |
| 39  | 0          | 0      |
| 40  | 0          | 0      |
| 41  | 0          | 0      |
| 42  | 0          | 0      |
| 43  | 0          | 0      |
| 44  | 0          | 0      |
| 45  | 0          | 0      |
| 46  | 0          | 0      |
| 47  | 0          | 0      |
| 48  | 0          | 0      |
| 49  | 0          | 0      |
| 50  | 0          | 0      |
| 51  | 0          | 0      |
| 52  | 0          | 0      |
| 53  | 0          | 0      |
| 54  | 0          | 0      |
| 55  | 0          | 0      |
| 56  | 0          | 0      |
| 57  | 0          | 0      |
| 58  | 0          | 0      |
| 59  | 0          | 0      |
| 60  | 0          | 0      |
| 61  | 0          | 0      |
| 62  | 0          | 0      |
| 63  | 0          | 0      |
| 64  | 0          | 0      |
| 65  | 0          | 0      |
| 66  | 0          | 0      |
| 67  | 0          | 0      |
| 68  | 0          | 0      |
| 69  | 0          | 0      |
| 70  | 0          | 0      |
| 71  | 0          | 0      |
| 72  | 0          | 0      |
| 73  | 0          | 0      |
| 74  | 0          | 0      |
| 75  | 0          | 0      |
| 76  | 0          | 0      |
| 77  | 0          | 0      |
| 78  | 0          | 0      |
| 79  | 0          | 0      |
| 80  | 0          | 0      |
| 81  | 0          | 0      |
| 82  | 0          | 0      |
| 83  | 0          | 0      |
| 84  | 0          | 0      |
| 85  | 0          | 0      |
| 86  | 0          | 0      |
| 87  | 0          | 0      |
| 88  | 0          | 0      |
| 89  | 0          | 0      |
| 90  | 0          | 0      |
| 91  | 0          | 0      |
| 92  | 0          | 0      |
| 93  | 0          | 0      |
| 94  | 0          | 0      |
| 95  | 0          | 0      |
| 96  | 0          | 0      |
| 97  | 0          | 0      |
| 98  | 0          | 0      |
| 99  | 0          | 0      |
| 100 | 0          | 0      |

| Query Match           | 2.28;            | Score 47;                   | DB 12;                      | Length 586;       |
|-----------------------|------------------|-----------------------------|-----------------------------|-------------------|
| Best local Similarity | 52.38;           | Prod. No. 0.56;             |                             |                   |
| Matches 104;          | Conservative     | 0;                          | Mismatches 95;              | Indels 0; Gaps 0; |
| QY 1853               | tttagagatcgtcttc | aaagaagtcacgcgttttag        | taocttaocttaaacgcggtttttt   | 1912              |
|                       |                  |                             |                             |                   |
| DB 257                | TTTTGCAATATTT    | TATAAAACTGAATCATCTCCAAAT    | TATTAGTATTGAAATATTTTTTCT    | 198               |
|                       |                  |                             |                             |                   |
| QY 1913               | gtttttaaagctcttc | caatcccaatcglagagttctt      | taatacaagagatatatttagagttt  | 1972              |
|                       |                  |                             |                             |                   |
| DB 197                | TTTTTAAATTTT     | CATGGATTTTTACCATAGGT        | TTTAAATAATCCACGATTAATTAATTT | 138               |
|                       |                  |                             |                             |                   |
| QY 1973               | ctgaaatcccaagat  | tatttttaaagcccatcttttttaggt | tgatgaatlaaasaattttta       | 2032              |
|                       |                  |                             |                             |                   |
| DB 137                | GAATATATTAAA     | AGTAATTAATAATTTTCTGTGAGAAA  | TTAAACATAAATACTTTTAA        | 78                |
|                       |                  |                             |                             |                   |
| QY 2033               | attaagcttttccag  | tgt 2051                    |                             |                   |
|                       |                  |                             |                             |                   |
| DB 77                 | ATAAATCTATG      | CTCCAGT 59                  |                             |                   |
|                       |                  |                             |                             |                   |

RESULT 8  
CNS010V2/C

|              |  |        |     |        |                 |
|--------------|--|--------|-----|--------|-----------------|
| CNS010UV2/7C | CNS010UV2  | 698 bp | DNA | linear | GSS 26-JUL-1999 |
| LOCUS        |  |        |     |        |                 |
| DEFINITION   | Drosophila melanogaster genome survey sequence T7 end of BAC<br>HACN05C08 of DrosBAC library from Drosophila melanogaster (fruit<br>fly), genomic survey sequence.                           |        |     |        |                 |
| ACCESSION    | AL039464   |        |     |        |                 |
| VERSION      | AL039464.1   |        |     |        |                 |
| KEYWORDS     | GI:5611075   |        |     |        |                 |
| SOURCE       | GSS.   |        |     |        |                 |
| ORGANISM     | fruit fly.<br>Drosophila melanogaster  |        |     |        |                 |
| REFERENCE    | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;<br>Neoptera; Endopterygota; Diptera; Brachycera;<br>Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.<br>1 (bases 1 to 698) |        |     |        |                 |

| AUTHORS |
|---------|
| TITLE   |
| JOURNAL |

Genoscope.  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) at  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaut at CSHG (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBel0BAC11.

FEATURES SOURCE

```

1. .698
/organism="Drosophila melanogaster"
/plasmid="pBel08AC11"
/db_xref="taxon:7221"
/clone_lib="DrosBAC"
/clone="BACN05C08"
/note="end : 177"

```

|            |       |      |      |       |            |
|------------|-------|------|------|-------|------------|
| BASE COUNT | 256 a | 88 c | 42 g | 127 t | 185 others |
| ORIGIN     |       |      |      |       |            |

Query Match 2.2%; Score 46.6; DB 12; Length 698;  
Best Local Similarity 35.4%; Pred. No. 0.73;  
Matches 85; Conservative 44; Mismatches 111; Indels 0

[illegible]

RESULT 9  
CNS04F5M

|   |
|---|
| CNS04E3M  |
| LOCUS   |
| CNS04E5M      976 bp    DNA     linear    GSS 21-MAY-2000       |
| Tetraodon nlgroviridis genome survey sequence T7 end of clone   |
| 103P02 of library G from Tetraodon nlgroviridis, genomic survey |
| DEFINITION  |

ACCESSION  
VERSION

VERSION  
AL28002.1 GI:8025084  
KEYWORDS  
GSS; genome survey sequence.  
SOURCE  
Tetraodon nigroviridis.  
ORGANISM  
Tetraodon nigroviridis

## ORGANISM

Actinopterygii; Neopreygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.  
 1 (bases 1 to 976)

## AUTHORS

|           |   |
|-----------|---|
| TITLE     | Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i> |
| JOURNAL   | Unpublished   |
| REFERENCE | 2 (bases 1 to 976)  |

## AUTHORS

TITLE Human gene number estimate provided by genome wide analysis using Saurin, W. and Weissenbach, J.









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 20:39:16 ; Search time 119.95 Seconds  
(without alignments)  
514.856 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 2951  
Sequence: 1 MSKLIRRVTVLALTSWASC.....ILSSDTLTSVPDTEHNVY 556

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID          | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 2951   | 100.0       | 556    | AAV35178    | Chlamydia pneumoni |
| 2          | 2951   | 100.0       | 556    | AA18820     | A 60 kda cysteine- |
| 3          | 2129.5 | 72.2        | 553    | AAV37788    | Chlamydia trachoma |
| 4          | 428    | 14.5        | 109    | AAK07399    | Chlamydia Crp gene |
| 5          | 130.5  | 4.4         | 1404   | AAV59600    | Drosophila Seriate |
| 6          | 130.5  | 4.4         | 1483   | ABG16336    | Novel human diagno |
| 7          | 129.5  | 4.4         | 757    | ABB70052    | Drosophila melanog |
| 8          | 129.5  | 4.4         | 1404   | AAK38304    | Sequence of a serr |
| 9          | 128.5  | 4.4         | 1404   | ABG61998    | Drosophila melanog |
| 10         | 124.5  | 4.2         | 1048   | 22 AAU07143 | Chicken CRIM1 prot |
| 11         | 122    | 4.1         | 2931   | 22 ABB68223 | Drosophila melanog |

|    |       |     |      |    |          |                    |
|----|-------|-----|------|----|----------|--------------------|
| 12 | 120.5 | 4.1 | 716  | 21 | AAV76005 | Rat Slit homologue |
| 13 | 120.5 | 4.1 | 716  | 22 | AAB55944 | Skin cell protein, |
| 14 | 119.5 | 4.0 | 2146 | 22 | ABB62317 | Drosophila melanog |
| 15 | 119   | 4.0 | 744  | 22 | ABB38426 | Drosophila melanog |
| 16 | 118.5 | 4.0 | 1529 | 21 | AAV76117 | Rat Slit homologue |
| 17 | 118.5 | 4.0 | 1529 | 22 | AAB56056 | Skin cell protein, |
| 18 | 117.5 | 4.0 | 1503 | 20 | AAV27142 | Human slit-2 matur |
| 19 | 117.5 | 4.0 | 1503 | 20 | AAW96701 | Slit-like protein  |
| 20 | 117.5 | 4.0 | 1525 | 20 | AAV17499 | Human slit-1 prote |
| 21 | 117.5 | 4.0 | 1529 | 20 | AAV27145 | Human slit-2 prote |
| 22 | 117.5 | 4.0 | 1529 | 20 | AAW96702 | Full length slit-1 |
| 23 | 117.5 | 4.0 | 1529 | 20 | AAB61239 | Human Slit-1 prote |
| 24 | 115.5 | 3.9 | 1036 | 22 | AAU07141 | Human CRIM1 protei |
| 25 | 115.5 | 3.9 | 1037 | 22 | AAU07142 | Mouse CRIM1 protei |
| 26 | 115.5 | 3.9 | 1810 | 17 | AAK94563 | Chicken cytotactin |
| 27 | 115.5 | 3.9 | 2312 | 22 | ABG06402 | Novel human diagno |
| 28 | 114.5 | 3.9 | 732  | 22 | AAB61140 | Human NOV10 protei |
| 29 | 114.5 | 3.9 | 1036 | 21 | AAV27776 | Human chordin rela |
| 30 | 114.5 | 3.9 | 1036 | 21 | AAV53034 | Human secreted pro |
| 31 | 114.5 | 3.9 | 1036 | 22 | AAU12242 | Human PRO4330 poly |
| 32 | 114.5 | 3.9 | 3478 | 22 | ABG04526 | Novel human diagno |
| 33 | 112.5 | 3.8 | 1358 | 17 | AAK99358 | Human restrictin.  |
| 34 | 112.5 | 3.8 | 1358 | 18 | AAK35743 | Human restrictin.  |
| 35 | 112.5 | 3.8 | 1358 | 18 | AAK18824 | B. sphaericus SLP. |
| 36 | 111.5 | 3.8 | 1252 | 16 | AAK80530 | Human tenascin-C.  |
| 37 | 111   | 3.8 | 2201 | 22 | AAK36935 | Slit protein. Xen  |
| 38 | 110.5 | 3.7 | 1530 | 21 | AAK28151 | Drosophila melanog |
| 39 | 110   | 3.7 | 2703 | 22 | ABG0266  | Human protein sequ |
| 40 | 109   | 3.7 | 413  | 22 | AAK92604 | Human cytotactin.  |
| 41 | 108.5 | 3.7 | 2199 | 17 | AAK94562 | Human ADAMTS-1 pro |
| 42 | 108   | 3.7 | 727  | 20 | AAW78435 | Human shear stress |
| 43 | 108   | 3.7 | 2705 | 22 | ABG90772 | Drosophila melanog |
| 44 | 108   | 3.7 | 3680 | 22 | ABW70878 | Human protein SEQ  |
| 45 | 107.5 | 3.6 | 1358 | 22 | AAW78544 |                    |

ALIGNMENTS

RESULT 1

ID AAV35178 standard; Protein; 556 AA.  
XX AAV35178;  
AC AAV35178;  
XX 13-SEP-1999 (first entry)  
DT Chlamydia pneumoniae transmembrane protein sequence.  
XX  
DE Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX Chlamydia pneumoniae.  
OS Chlamydia pneumoniae.  
XX  
PN WO9927105-A2.  
XX  
PD 03-JUN-1999.  
XX  
PF 20-NOV-1998; 98WO-IB01890.  
XX  
PR 04-NOV-1998; 98US-0107078.  
PR 21-NOV-1997; 97FR-0014673.  
XX (GIST ) GENSET.  
XX Griffais R;  
XX WPI; 1999-357842/30.  
XX Genome sequence of Chlamydia pneumoniae  
XX Page 1037-1038; Disclosure; 1912pp; English.  
PS

XX AAY34584-Y35879 represent the proteins encoded by all the open reading  
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.  
XX  
SQ Sequence 556 AA;

Query Match 100.0%; Score 2951; DB 20; Length 556;  
Best Local Similarity 100.0%; Pred. No. 4.7e-233;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVVTVLALTSMAFCASGGIEAAVASLITKIVASAEKPPAPVMTAKKVLVR 60  
DB 1 MSKLIRRVVTVLALTSMAFCASGGIEAAVASLITKIVASAEKPPAPVMTAKKVLVR 60  
QY 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCGRLYSVKVNDDCNVEICOSVPE 120  
DB 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCGRLYSVKVNDDCNVEICOSVPE 120  
QY 121 YATVGSPPYPIELAIAGKDKCDVVDVITQOLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180  
DB 121 YATVGSPPYPIELAIAGKDKCDVVDVITQOLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
DB 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNPPVPGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300  
DB 241 EVVNTGSAIARNVTVDNPPVPGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300  
QY 301 NVATVTCGSHKCSANVTTVVNEPCQVNIAGADWSYCKPVEYSISVSNPGLVLDHVV 360  
DB 301 NVATVTCGSHKCSANVTTVVNEPCQVNIAGADWSYCKPVEYSISVSNPGLVLDHVV 360  
QY 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKVKAQVPGFTNOVAVTSE 420  
DB 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKVKAQVPGFTNOVAVTSE 420  
QY 421 SNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNRGSAEDTNVSLIL 480  
DB 421 SNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNRGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDAIPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540  
DB 481 KFSKELOPIASSGPTKGTISGNTVVFDAIPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHY 556  
DB 541 DTLTSPVSDTENTHY 556

## RESULT 2

AAB18820  
ID AAB18820 standard; Protein; 556 AA.  
XX  
AC AAB18820;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE A 60 kDa cysteine-rich membrane protein of Chlamydia pneumoniae.  
XX  
KW Cysteine-rich membrane protein; Chlamydia infection; bronchitis;  
KW community acquired pneumonia; upper respiratory tract infection; vaccine;

KW sinusitis.  
XX Chlamydia pneumoniae.  
XX WO200055326-A1.  
XX 21-SEP-2000.  
XX  
XX 09-MAR-2000; 2000WO-CA00240.  
XX 12-MAR-1999; 99US-0123966.  
XX (AVET ) AVENTIS PASTEUR LTD.  
XX  
XX Murdin AD, Omen RP, Wang J, Dunn P;  
XX WPI; 2000-618918/59.  
XX N-PSDB; AAA75901.  
XX  
XX New polynucleotides encoding a 60kDa cysteine-rich membrane protein  
XX from Chlamydia, useful as a vaccine for preventing and treating  
XX Chlamydia infection in mammals -  
XX  
XX Claim 16; Fig 1; 77pp; English.  
XX  
XX The present sequence represents a Chlamydia 60 kDa cysteine-rich membrane  
XX protein. The membrane-rich polynucleotide and polypeptide are useful  
XX for preventing or treating Chlamydia infection, such as community  
XX pneumonia, upper respiratory tract infections, bronchitis and sinusitis.  
XX They are also useful for diagnosing Chlamydia infection by assaying a  
XX body fluid of a mammal. The polypeptide is useful for vaccine  
XX production.  
XX  
SQ Sequence 556 AA;

Query Match 100.0%; Score 2951; DB 21; Length 556;  
Best Local Similarity 100.0%; Pred. No. 4.7e-233;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVVTVLALTSMAFCASGGIEAAVASLITKIVASAEKPPAPVMTAKKVLVR 60  
DB 1 MSKLIRRVVTVLALTSMAFCASGGIEAAVASLITKIVASAEKPPAPVMTAKKVLVR 60  
QY 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCGRLYSVKVNDDCNVEICOSVPE 120  
DB 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCGRLYSVKVNDDCNVEICOSVPE 120  
QY 121 YATVGSPPYPIELAIAGKDKCDVVDVITQOLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180  
DB 121 YATVGSPPYPIELAIAGKDKCDVVDVITQOLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
DB 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNPPVPGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300  
DB 241 EVVNTGSAIARNVTVDNPPVPGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300  
QY 301 NVATVTCGSHKCSANVTTVVNEPCQVNIAGADWSYCKPVEYSISVSNPGLVLDHVV 360  
DB 301 NVATVTCGSHKCSANVTTVVNEPCQVNIAGADWSYCKPVEYSISVSNPGLVLDHVV 360  
QY 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKVKAQVPGFTNOVAVTSE 420  
DB 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKVKAQVPGFTNOVAVTSE 420  
QY 421 SNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNRGSAEDTNVSLIL 480  
DB 421 SNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNRGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDAIPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540

[illegible]



```
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS80523.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20; SEQ ID NO 46695; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX mapping and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1483 AA;
XX
XX Query Match 4.4%; Score 130.5; DB 22; Length 1483;
XX Best Local Similarity 19.3%; Pred. NO. 0.18;
XX Matches 118; Conservative 90; Mismatches 217; Indels 185; Gaps 29;
XX
XX 35 ITKIVASAEKPAV-----PMTAKKVR-----LVRRNKQPVQKSRGAFCDKE--- 78
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 626 vsqvgscetlhpdrfnrreiqgafrlnrlntllnnq-ikripsgafedlenlk 684
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 79 FYPCBEGRCQPVDAQOESCYGRLYSVKVNDD---CNVEI---CQSVPEYATVGS----- 126
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 685 yilykneiqsdrqafglaeqrlrldstlhcdeallwladilktyaesgnaqaaai 744
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 127 -PYPIEILAIKGDVYVITQOLCEAEFVSDDPTTSDGKLVKIDRLGAGDKCKI 185
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 745 ceyprri-----qgrsvatitpeelnrcerpritsepqdaavtsqntvyfctraegnpkpei 800
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 186 TWVW--KPLKEGCGCTTAATVACAPELRSYTKGQPAICIKQEGPCACALRCP--VCYKIEV 242
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 801 -Iwlrnkslerhaaadtslssrvdl-----tcipnkg---slllcpqtccgpyk 846
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 243 VNTGSAIARNVTVDNPPVPGVSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRPQRIITNV 302
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 847 aktalaaavrvtegsptadhy-----mrpgsephslsltc---rlalpmi 898
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 303 -----ATVT-----YGGGHKCSANVTTVVNEPCVQVNIAGDSWYVC---KPVEYSISVS 349
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 889 qshlqatltcwlavfltgisvasnvtsl-----ftcfclkgtkvklqv 933
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 350 NPGDL-----VLHD--VVIQDTLPSGVTVLEA-----PGGEICCNKVVVRIKEMCPGET 396
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 934 nelsmktdsrlnliddgtlmiqntqetdqqiyqcmaknvagevktqev-----t 982
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 397 LOFKLVVKAQVPGREFT-NOVAVTSESNGCTCTSCAETTHWKGLAATHMCVLDTNDP--- 452
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 983 lry-----fgspgsflgnrvctsvsvctatslqpqntevlgvesvtlecatghpppri 1037
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 453 -----ICVGENTVYRICVTNTRGSAEDTNVSLILKFSK 484
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1038 swtrgdrtpdpvdrvnitpsgglyiqnvvgdsgdeyacsatnnidsvhatafiiv---q 1094
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 485 ELQPIASSGPTKGTIGSNTVVFD-----ALPKLGSKSESVFSTLKIAPGDAR 533
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1095 alpqtvtvpqdrvvlegqtvdvqceakgnpppviawtkggsqlsvd-----r 1141
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 534 GEATLSSDTL 543
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1142 rhlvissgtl 1151
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
ABB70052
ID ABB70052 standard; Protein; 757 AA.
XX
XX ABB70052;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 36948.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX N-PSDB; ABL14155.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions
XX
XX Disclosure; SEQ ID NO 36948; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 757 AA;
XX
XX Query Match 4.4%; Score 129.5; DB 22; Length 757;
```





Tue May 28 08:55:16 2002

```
DE Drosophila melanogaster polypeptide SEQ ID NO 12786.
XX Drosophila: developmental biology; cell signalling; insecticide;
XX Pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL06101.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Disclosure; SEQ ID NO 12786; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA
XX sequences (ABU01840-ABU16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 1404 AA;
XX
XX Query Match 4.4%; Score 128.5; DB 22; Length 1404;
XX Best Local Similarity 20.5%; Pred. No. 0.24; Mismatches 217; Gaps 32;
XX Matches 124; Conservative 57;
QY 14 LTSMSASFASGGTEAAVAESLTIKIVASAEETKPAVPMTAKKRVLRNKPVEQKSRGA 73
Db 543 lttataigsnlsstallaaltsavastalaigpc-inakecr-----nqp-----gs 590
QY 74 FCDKEFYPCBEGRCQPVEAQOQESCYGRLYSVKVNDCNVCICQSVPEYATVGSYPYIEIL 133
Db 591 fa----cicqg-----wggvtcaenlddc-vdgcrn----- 617
QY 134 AIGKKCDVDVYITQQLPCEAFFVSSDPE-----TPTSDG-----KLVWK-----IDRLG- 178
Db 618 ---gatkdivndyracasgfgkgrdcetdidecatpcrnnnggcvdmvgkfcncicply 674
QY 179 AGDKC---KITVMVKPLKEGSCCFTA--ATVCACPELRSYTKCGQ-PAIC----- 221
Db 675 sgslceakentpcpclegchclntpegyychpcpdragkhceqlrplcsqppcneqgfa 734
QY 222 -----IKQEGPDC-----ACLRCPV---CYKIEWVNTGSAIA 250
Db 735 nvslatsattttttttrkrmakpsglpcsghgscemsdvgtfc-kchvghtgfce 793
QY 251 RNVTVDNVPP-----DG-----YSHASGORVLS-----FNLGDMRPG-- 282
Db 794 hmlncspnrcnggicldgdgdfccemsgwtgkrseratgcyagqcngggtcmppgap 853
QY 283 DKKV-----FTVEFCPQRRGQITNVATVTCYGGHKCSANVTTVNPEPCVQVNISGA 333
```

```
RESULT 10
AAU07143
ID AAU07143 standard; Protein; 1048 AA.
XX
XX AAU07143;
XX AC
XX DT
XX 24-OCT-2001 (first entry)
XX DE
XX Chicken CRIM1 protein.
XX CRIM-1; Chicken; human chromosome 2p21-16.3; ophthalmological;
XX neuroprotective; renal; osteopathic; dental; vulnary; immunogen;
XX antibody; gene therapy; neurodegenerative disease; eye disorder;
XX cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
XX tooth abnormality; wound; S52.
XX OS
XX Gallus gallus.
XX
XX Key Location/Qualifiers
XX Region 212..219
XX /note= "conserved N-terminal motif"
XX Region 348..402
XX /label= CR_1
XX /note= "Cysteine rich repeat"
XX Region 415..468
XX /label= CR_2
XX /note= "Cysteine rich repeat"
XX Region 620..674
XX /label= CR_3
XX /note= "Cysteine rich repeat"
XX Region 691..746
XX /label= CR_4
XX /note= "Cysteine rich repeat"
XX Region 765..820
XX /label= CR_5
XX /note= "Cysteine rich repeat"
XX Region 831..885
XX /label= CR_6
XX /note= "Cysteine rich repeat"
XX
XX WO200138519-A1.
XX
XX 31-MAY-2001.
XX
XX 24-NOV-2000; 2000WO-AU01435.
XX
XX 26-NOV-1999; 99AU-0004348.
XX
XX (UYQU ) UNIV QUEENSLAND.
XX
XX Little M, Yamada T, Holmes G, Georgas K, Kolle G, Wilkinson L;
XX WPI; 2001-343951/36.
XX
XX DB
```

```
Db 854 dkalqphrcapgwglfcae-----aidqcrqpcchnggt-----ce-----sga 894
QY 334 DW-SYVCKPEVYSISVSNPGDLVLHDVWTDLPVSGVTVLEAPGGEICCNKVVMRIKEMC 392
Db 895 gwfrvcv-----agfgsgpdcrlinvnecspqpcggatcidgiggygc-----ic 939
QY 393 PGETLQFLVVKVKAQVGRFTNQAVTSESCGCTCTSCAET-----TTHWKGLAATH 443
Db 940 p-----pgrhglrceillsdpsacqnasntispytalnrsgnwidialtg 985
QY 444 MCVLDTNDPICVGVTVYRICVTNRGSAEDTNVSLILKFSKELOPIASSGPTKTIGTSNT 503
Db 986 rteddencnaccven-----gtsrotnlwcgipncykvdplskssnlsnsvckqkhe 1035
QY 504 VVFDAL 509
Db 1036 vcvpal 1041
```



Tue May 28 08:55:16 2002

CC AAY76119 are proteins with an N-terminal signal sequence, indicating  
CC that they are secreted. Sequences AAY75986-f75989, AAY76061-Y76071,  
CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more  
CC putative transmembrane domains.  
XX  
SQ Sequence 716 AA;

Query Match 4.1%; Score 120.5; DB 21; Length 716;  
Best Local Similarity 20.4%; Pred. No. 0.42; Indels 191; Gaps 24;  
Matches 92; Conservative 42; Mismatches 126; Indels 191; Gaps 24;

QY 72 GAFCDKEFPCEGRGQ-----PVEAQQESQ-----YGRLY-----SVKVNDD 109  
Db 313 gehcdidfdcdqkngahctdavnngytcvcepgysglfcefppmvflrtspcdnfd 372  
QY 110 C-----NVEICOSPEYATVGSPPYPIEILAIKDKDC-----VDVVITQQLPCEAE 154  
Db 373 cqnagaqcilrvnepicqclpgy-----lgek-ceklsvsllivnkesylqip 418  
QY 155 FVSSDPET-----TPTSDGKLWKIDRLGAGDKCKITVMVKPLKFGCCFTTATVCACP 207  
Db 419 sakvrpqtntlqiatdedsqillyk-----gdkdhav-----esiegi----- 458  
QY 208 ELRGYTKGQPAICIKQEGPDCACLRCPVCYKIEVWTGS-----AIAKNVTVDNP 258  
Db 459 -rasydtgshpasal-----ysvetindgnfhiwelltdsslsldvsg 502  
QY 259 VPDGYSHASGQVLSFN-----LGDMPGDKKVFVFEFCQRRGQITNVATVYTCGGHKCS 314  
Db 503 spkiitniskstlnfdsplyvggm-pgknnavaslrqapqng-----tsfhgci 551  
QY 315 ANVTWVNEPCVQVNIAGDMSYVCKPVEYSISVNSPGDLVLDVVIODTLPSGVTVLEA 374  
Db 552 rh-----lyinselqdfkrkv--pmgtgi-----l 573  
QY 375 PGGEICCNKVVWRKEMCPGETLQFKLVVKAQVPGRFTNQAVTSESCGTCTCAETT 434  
Db 574 pgcepchkkv-----canghtc-----psqsqgfcece----- 603  
QY 435 HWKGLAATHMCVLDNDPICVGVNTVYRICV 465  
Db 604 gwmq-----plcdqrtdnp-clgnkvcvhtcl 629

RESULT 13  
AAB55944  
ID AAB55944 standard; Protein; 716 AA.  
XX  
AC AAB55944;  
XX  
DT 08-MAR-2001 (first entry)  
XX  
DE Skin cell protein, SEQ ID NO: 183.  
XX  
KW Rat; skin cell; cytostatic; antinflammatory; anti-HIV;  
KW neutropic; neuroprotective; vulnerary; immunomodulatory; vaccine;  
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;  
KW inflammation; neurological disease.  
XX  
OS Rattus sp.  
XX  
PN WO200069884-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 15-MAY-2000; 2000WO-NZ00075.  
XX  
PR 14-MAY-1999; 99US-0312283.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;

Db 1473 csqgkneysitpdpntclg-----ctstdprcatnqnqltttnicsqipyte 1521  
QY 368 GVTVEAPGGEI--C-----CNKVVWRKEMCPGETLQFKLVYKAQVPGRFTNQ 414  
Db 1522 cvtdidsqntirgcmsslsgddfyecitgdgkicetctgerngisvfpad--rrkcyg 1579  
QY 415 VAVTSENGCTCTCAETTHWKLAAATHMCVLDNDPICV-----IASSGPTK 496  
Db 1580 cnstdpncaspatlest-----vcpiysqdescvttllnglthrgcscslltcs 1629  
QY 456 --GENTVYRICVTRNGSAEDTNVSLILKFSKELQP-----IASSGPTK 496  
Db 1630 dpsdrtcrvc-----ssadgcntinlekigedfgpgawdvpieclvcsgtcdassggtl 1685  
QY 497 GTISGN-----TWVFD--ALPKLGSKEV 518  
Db 1686 tkcsydnvcvtvdsdgsavtrgcesv 1713

RESULT 12  
AAY76005  
ID AAY76005 standard; Protein; 716 AA.  
XX  
AC AAY76005;  
XX  
DT 27-MAR-2000 (first entry)  
XX  
DE Rat Slit homologue, SEQ ID NO:183.  
XX  
KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
KW secreted; transmembrane; inflammation; cancer; neurological disease;  
KW angiogenesis; tumour vasculature; growth disorder;  
KW developmental disorder; skin wound; hair follicle disorder;  
KW anti-inflammatory; cytostatic; neuroprotective; vulnerary.  
XX  
OS Rattus sp.  
XX  
PN WO9955865-A1.  
XX  
PD 04-NOV-1999.  
XX  
PF 29-APR-1999; 99WO-NZ00051.  
XX  
PR 29-APR-1998; 98US-0069726.  
XX  
PR 09-NOV-1998; 98US-0188930.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;  
XX  
DE WPI: 2000-072177/06.  
XX  
DR N-PSDB; AA261669.  
XX  
PT Novel polynucleotides useful for the treatment of various conditions  
XX including wounds and cancer -  
XX  
PS Claim 4; Page 125-126; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat  
XX dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
XX and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
XX cells. Polypeptides of the invention may be used to treat inflammation,  
XX cancer and neurological diseases. The proteins may be used to stimulate  
XX the growth and motility of keratinocytes, to inhibit the growth of  
XX cancer cells, to modulate angiogenesis and tumour vascularisation, to  
XX modulate skin inflammation, to modulate epithelial cell growth and to  
XX inhibit binding of HIV-1 to leukocytes. The invention may also be used  
XX to treat growth and developmental defects, skin wounds and hair follicle  
XX disorders. Sequences AAY75942-Y76123 represent polypeptides encoded  
XX by cDNA sequences derived from several mouse, rat or human skin cell  
XX types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and







Tue May 28 08:55:17 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 20:42:06 : Search time 54.34 Seconds  
(without alignments)  
249.920 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 2951  
Sequence: 1 MSKLIRRVTVLALTMASC.....ILSSDTLTSPVSDPTENTHYV 556

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues  
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Match | Length | DB ID             | Description       |
|------------|-------|---------|-------|--------|-------------------|-------------------|
| 1          | 435   | 14.7    | 109   | 1      | US-08-477-270-20  | Sequence 20, Appl |
| 2          | 141   | 4.8     | 25    | 4      | US-09-025-596-38  | Sequence 38, Appl |
| 3          | 130.5 | 4.4     | 1404  | 2      | US-08-400-159-2   | Sequence 2, Appl  |
| 4          | 130.5 | 4.4     | 1404  | 3      | US-08-611-729A-2  | Sequence 2, Appl  |
| 5          | 120.5 | 4.1     | 771   | 4      | US-09-188-930-183 | Sequence 183, App |
| 6          | 119   | 4.0     | 20    | 4      | US-09-025-596-41  | Sequence 41, Appl |
| 7          | 117.5 | 4.0     | 1525  | 3      | US-09-191-647-2   | Sequence 2, Appl  |
| 8          | 117.5 | 4.0     | 1525  | 4      | US-09-540-245A-2  | Sequence 2, Appl  |
| 9          | 117.5 | 4.0     | 1525  | 4      | US-09-540-153-2   | Sequence 4, Appl  |
| 10         | 115.5 | 3.9     | 1810  | 5      | PCT-US95-11684-4  | Sequence 4, Appl  |
| 11         | 112.5 | 3.8     | 1358  | 1      | US-08-404-665-4   | Sequence 4, Appl  |
| 12         | 112.5 | 3.8     | 1358  | 1      | US-08-404-671-4   | Sequence 4, Appl  |
| 13         | 112.5 | 3.8     | 1358  | 1      | US-08-404-781-4   | Sequence 4, Appl  |
| 14         | 111.5 | 3.8     | 1222  | 2      | US-08-682-517-15  | Sequence 15, Appl |
| 15         | 111.5 | 3.8     | 1252  | 2      | US-08-682-517-9   | Sequence 9, Appl  |
| 16         | 108.5 | 3.7     | 2199  | 5      | PCT-US95-11684-2  | Sequence 2, Appl  |
| 17         | 106.5 | 3.6     | 4654  | 4      | US-08-476-515A-84 | Sequence 84, Appl |
| 18         | 106.5 | 3.6     | 4655  | 4      | US-08-652-877-84  | Sequence 84, Appl |
| 19         | 106.5 | 3.6     | 4655  | 4      | US-08-652-877-86  | Sequence 86, Appl |
| 20         | 106.5 | 3.6     | 4655  | 4      | US-08-652-877-88  | Sequence 88, Appl |
| 21         | 106.5 | 3.6     | 4655  | 4      | US-08-652-877-90  | Sequence 90, Appl |
| 22         | 106   | 3.6     | 19    | 4      | US-09-025-596-39  | Sequence 39, Appl |
| 23         | 105.5 | 3.6     | 2813  | 3      | US-08-896-449A-2  | Sequence 2, Appl  |
| 24         | 105.5 | 3.6     | 2813  | 3      | US-09-132-652-2   | Sequence 2, Appl  |
| 25         | 105   | 3.6     | 2556  | 1      | US-08-083-590A-20 | Sequence 20, Appl |
| 26         | 105   | 3.6     | 2556  | 1      | US-08-532-384-20  | Sequence 20, Appl |
| 27         | 104   | 3.5     | 886   | 3      | US-09-110-116-3   | Sequence 3, Appl  |

|    |       |     |      |   |                   |                   |
|----|-------|-----|------|---|-------------------|-------------------|
| 28 | 104   | 3.5 | 2703 | 1 | US-08-185-432-19  | Sequence 19, Appl |
| 29 | 103.5 | 3.5 | 2471 | 1 | US-08-185-432-16  | Sequence 16, Appl |
| 30 | 103.5 | 3.5 | 2471 | 1 | US-08-083-590A-19 | Sequence 19, Appl |
| 31 | 103.5 | 3.5 | 2471 | 3 | US-08-532-384-19  | Sequence 19, Appl |
| 32 | 102   | 3.5 | 2523 | 1 | US-08-185-432-18  | Sequence 18, Appl |
| 33 | 101.5 | 3.4 | 1480 | 3 | US-09-191-647-7   | Sequence 7, Appl  |
| 34 | 101.5 | 3.4 | 1480 | 4 | US-09-540-245A-7  | Sequence 7, Appl  |
| 35 | 101.5 | 3.4 | 1480 | 4 | US-09-540-153-7   | Sequence 7, Appl  |
| 36 | 101.5 | 3.4 | 1480 | 5 | PCT-US91-09035-2  | Sequence 2, Appl  |
| 37 | 100   | 3.4 | 2556 | 1 | US-08-185-432-17  | Sequence 17, Appl |
| 38 | 99    | 3.4 | 1940 | 2 | US-08-644-271-30  | Sequence 30, Appl |
| 39 | 98.5  | 3.3 | 2647 | 2 | US-08-583-562B-8  | Sequence 8, Appl  |
| 40 | 98.5  | 3.3 | 2647 | 2 | US-08-779-113-8   | Sequence 10, Appl |
| 41 | 97.5  | 3.3 | 1193 | 3 | US-08-400-159-10  | Sequence 10, Appl |
| 42 | 97.5  | 3.3 | 1193 | 3 | US-08-611-729A-10 | Sequence 5, Appl  |
| 43 | 97.5  | 3.3 | 1480 | 4 | US-09-182-024A-5  | Sequence 2, Appl  |
| 44 | 97.5  | 3.3 | 2050 | 2 | US-08-347-594A-2  | Sequence 2, Appl  |
| 45 | 96.5  | 3.3 | 1833 | 3 | US-08-479-722B-2  | Sequence 2, Appl  |

ALIGNMENTS

RESULT 1  
US-08-477-270-20  
: Sequence 20, Application US/08477270  
: Patent No. 5629158  
: GENERAL INFORMATION:  
: APPLICANT: UHLEN, Mathias  
: TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL  
: TITLE OF INVENTION: CONDITIONS  
: NUMBER OF SEQUENCES: 30  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 1800 Diagonal Road, Suite 500  
: CITY: Alexandria  
: STATE: Virginia  
: COUNTRY: USA  
: ZIP: 22313-0299  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/477,270  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: PRIOR APPLICATION NUMBER: US/08/261,010  
: FILING DATE:  
: APPLICATION NUMBER: US 07/781,157  
: FILING DATE: 07-NOV-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: BENT, Stephen A.  
: REGISTRATION NUMBER: 29,768  
: REFERENCE/DOCKET NUMBER: 16787/153 DFBC  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703)836-9300  
: TELEFAX: (703)683-4109  
: TELEX: 899149  
: INFORMATION FOR SEQ ID NO: 20:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 109 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-477-270-20

Query Match 14.7%; Score 435; DB 1; Length 109;  
Best Local Similarity 72.5%; Pred. No. 6.7e-33;  
Matches 79; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

[illegible]

```

RESULT      2
US-09-025-596-38
; Sequence 38, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025.596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-38

```

```

Query Match          4.8%;   Score 141;   DB 4;   Length 35;
Best Local Similarity 100.0%;   Pred. No. 1.2e-06;
Matches 25; Conservative 0; Mismatches 0; Indels
QY 60 RNNQVPEQKSRGAFCDKREFYPCEE 84
      |||
db 1 RNNQVPEQKSRGAFCDKREFYPCEE 25

```

RESULT 3  
 US-840-159-2  
 ; Sequence 2, Application US/08400159  
 ; Patent No. 5869282  
 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ish-Horowicz, David  
 ; APPLICANT: Henrique, Domingos M.P.  
 ; APPLICANT: Lewis, Julian H.  
 ; APPLICANT: Myat, Anna M.  
 ; APPLICANT: Fleming, Robert J.  
 ; APPLICANT: Artavanis-tsakonas, Spyridon  
 ; APPLICANT: Mann, Robert S.  
 ; APPLICANT: Gray, Grace E.  
 ;  
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE  
 ; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON  
 ;  
 ; NUMBER OF SEQUENCES: 20  
 ;  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ;  
 ; ZIP: 10036-2711  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ;  
 ; CURRENT APPLICATION DATA:

```

: APPLICATION NUMBER: US/08/400,159
: FILING DATE: 07-MAR-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 7326-029
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 859-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1404 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-400-159-2

```

|                       |              |  |                |             |         |
|-----------------------|--------------|--|----------------|-------------|---------|
| Query Match           | 4.4%         | Score 130.5  | DB 2           | Length 1404 |         |
| Best Local Similarity | 20.5%        | Pred. No. 0.0041   |                |             |         |
| Matches 124           | Conservative | 57   | Mismatches 208 | Indels 217  | Gaps 32 |
| QY                    | 14           | LTSWASCFASGGIEAAVAESLIITKIVASAETKPAPVPMWAKVRLVRRNKPQPVOKSGRA   | 73             |             |         |
| DB                    | 543          | LTITATAIIGSNLSSTALLAAALTSAVASTSLAIGPC-INAKECR-----NQP-----GS   | 590            |             |         |
| QY                    | 74           | FCDKEFYPCBEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSVPEYATVGSPPYTEIL   | 133            |             |         |
| DB                    | 591          | FA-----CICKEG-----WGVTCENLDDC-VGOCRN-----                      | 617            |             |         |
| QY                    | 134          | AIGKDKCDVVIIQOOLPCEAEFYSSDPE-----TPTSDG-----KLWVK-----IDRLG-   | 178            |             |         |
| DB                    | 618          | --GATCIDLNDVYRCASAGFTGRDCEITDIDECATSPCRNGBCGCDVMVGKFCICPLGY    | 674            |             |         |
| QY                    | 179          | AGDKC---KITVMVKPLKEGCCFTA--ATVCACPELRSYTKCGQ-PAIC-----         | 221            |             |         |
| DB                    | 675          | SGSLCEAEAKENCPTSPCLEGHCLNTPBGYYCHCPDPORAGKHCBOLRPLCSQPCCNEGCFA | 734            |             |         |
| QY                    | 222          | -----IKQGGPDC-----ACLRCPV---CYKIEVNTVNGSAIA                    | 250            |             |         |
| DB                    | 735          | NVSLATSATTTTTTTTATTTRKMAKPSGLPCSGHSGSCMSDVGTFEC-KCHVGHTGTGFC   | 793            |             |         |
| QY                    | 251          | RNVTDNPNP-----DG-----YSHASGQRVLS-----FNLGDMRPG--               | 282            |             |         |
| DB                    | 794          | HNLNECSPNCRNGICLDGDGDFCECHSGWTKRCSEERATGCYAGOCONGGTCTMPGAP     | 853            |             |         |
| QY                    | 283          | DKKV-----FTVEFCPQRGQITNVATVTCYCGHKCSANVTTVNEPCVOVNISGA         | 333            |             |         |
| DB                    | 854          | DKALOPHCAPGWTGLFCAE-----AIDCRQCPCHNGGT-----CE-----SGA          | 894            |             |         |
| QY                    | 334          | DW-SYVCKPVEYSISVSNPGLVLHDVLIQDTLPSGVTVLEAPGGEICCNKVVWRKEMC     | 392            |             |         |
| DB                    | 895          | GWFRVCV-----AQGFSGPDCRLNVNECSQPQCGGATCIDIGGYSC-----IC          | 939            |             |         |
| QY                    | 393          | PGETLQFLKLVAAQVPGRTNVOAVTSESNCGTCTCAET-----TTHWKGIAATH         | 443            |             |         |
| DB                    | 940          | P-----PGRHGLRCCELLSDPKSACQNASNTISPYTALNRSQNWLDIALTC            | 995            |             |         |
| QY                    | 444          | MCVLDNDPICVGVNTVYRICVTNRGSADETNVSLILKESKELOPTASSGPTKGTISGWT    | 503            |             |         |
| DB                    | 986          | RTEDDENCNACVCN-----GTSRCTNLWCGLPNCYKVDPLSKSSNLGCVCKOHE         | 1035           |             |         |
| QY                    | 504          | VVFDAI   | 509            |             |         |
| DB                    | 1036         | VCVPAL   | 1041           |             |         |

RESULT 4  
US-08-611-729A-2  
; Sequence 2, Application US/08611729A  
; Patent No. 6004924



GENERAL INFORMATION:  
 APPLICANT: Ish-Horowitz, David  
 APPLICANT: Henrique, Domingos M.P.  
 APPLICANT: Lewis, Julian H.  
 APPLICANT: Myat, Anna M.  
 APPLICANT: Fleming, Robert J.  
 APPLICANT: Artavanis-Tsakonas, Spyridon  
 APPLICANT: Mann, Robert S.  
 APPLICANT: Gray, Grace E.  
 TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE  
 TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/611,729A  
 FILING DATE: 06-MAR-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Misrock, S. Leslie  
 REGISTRATION NUMBER: 18,872  
 REFERENCE/DOCKET NUMBER: 7326-037  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-9741/8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1404 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08-611-729A-2

Query Match 4.4%; Score 130.5; DB 3; Length 1404;  
 Best Local Similarity 20.5%; Pred. No. 0.0041;  
 Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;  
 QY 14 LFSMASCSPASGGIEAAVASELITKIVASAEKTPAPVPTAKKVLVRRNKQPEQKSRGA 73  
 DB 543 LTTATAIGSLSTALLAALTSVAVSTSLAIGPC-INAKCR-----NQP-----GS 590  
 QY 74 FCDKEFYPCCEGRCPVEAQESCYGRYSVKVNDNCNVICQSPVEATVGSPPYIEL 133  
 DB 591 FA-----CICKEG-----WGGVTCAEMLDDC-VGQCRN----- 617  
 QY 134 AIGKDCVDVITQQLPCAEEFVSSDPE-----TTPSTDG-----KLVWK---IDRLG- 178  
 DB 618 ---GATCIDLVNDYRCACASGFTGRCDTIDECATSPCRNGECVDVMVGKENCICPLGY 674  
 QY 179 AGDKC---KITVVMVKPLKECCCFPA--ATVCACPELRSYTKCQ-PAIC----- 221  
 DB 675 SGLCEAEAKENCTPSCLEGHCLNTPEGYCHCPDPRAGKHCBQLRPLCSQPPCNGGCPA 734  
 QY 222 -----IKOEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250  
 DB 735 NVSLATSATTTTTTTTATTTRKMAKPSGLPCSHGSCMSDVGTFC-KCHVGHGTGTCF 793  
 QY 251 RNVTVDNVPP-----DG-----YSHASQORYLS-----ENLGDMPRG-- 282  
 DB 794 HNLNCSNPNCRNGICLDGDDFTCBMCSGWTCKRCSEKATGCGYAGCQCGQNGGTCMPGAP 853

QY 283 DKKV-----FTVEFCPORRGQITNVAITVTCYGGHKCSANVTTVVNEPCVQVNISGA 333  
 DB 854 DKALQPHCRCAPGWTGLFCAE-----AIDQCRGQPCNGGT-----CE-----SGA 894  
 QY 334 DW-SYVCKPVEYSISNPGDLVLDHVVQIDTLPSPGVTVLEAPGGEICCNKVVRHKEWC 392  
 DB 895 GWFRCVC-----AQFSGPDCRINYNECSPOPCGGATCIDGIGGYS-----IC 939  
 QY 393 PGETLQFKLVVKAQVPGREFTNQAVTSESNGCTCTSCAET-----TTHWKGLAATH 443  
 DB 940 P-----PGRHGLRCEILLSDPKSACONASNTISPYTALNRSONWLDIALTG 985  
 QY 444 MCVLDTPICVGTIVYRICVTNRGSAEDTNVSLTKFSEKELQPIASSGPTKGTISGNT 503  
 DB 986 RTEDDENCNACVCN-----GTSRCTNLWCGLPNCYKVDPLSKSNLSGVCCKQHE 1035  
 QY 504 VVFDAL 509  
 DB 1036 VCVPAL 1041  
 RESULT 5  
 US-09-188-930-183  
 ; Sequence 183, Application US/09188930A  
 ; Patent No. 6150502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew  
 ; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murlison, James Greg  
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; TITLE OF INVENTION: and Methods For Their Use  
 ; FILE REFERENCE: 11000.1011c1  
 ; CURRENT APPLICATION NUMBER: US/09/188.930A  
 ; CURRENT FILING DATE: 1998-11-09  
 ; NUMBER OF SEQ ID NOS: 348  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 183  
 ; LENGTH: 771  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (717)...(717)  
 ; US-09-188-930-183

Query Match 4.1%; Score 120.5; DB 4; Length 771;  
 Best Local Similarity 20.4%; Pred. No. 0.014;  
 Matches 92; Conservative 42; Mismatches 126; Indels 191; Gaps 24;  
 QY 72 GAFCDKEFYPCCEGRCP-----PVEAQOESC---YGRLY-----SVKYNDD 109  
 DB 313 GEHCIDIDFDDQDNKCKNGAHCCTDAVNGYTCVCPGEGYGLFCFSPPMVFLRTSPCNFD 372  
 QY 110 C-----NVICQSPVEATVGSPPYIEILAIAGCKDQ-----VDVITQQLPCEAE 154  
 DB 373 CONGAQCIIRVNEPICQCLPGY-----LGEK-CEKLVSVSILVKNESYLQIP 418  
 QY 155 FVSSDPEI-----TPTSDGKLVWKIDRLGAGDKCKITVWVKPLKECCCFATATVCACP 207  
 DB 419 SAKVRPQTNITLQIATDEDSGLLYK-----GDKDHIAV---BSIEGI----- 458  
 QY 208 ELRSYTKCQOPACICIKQEGPCACLRCPVCYKIEVVNTGS-----ATARNVTVDNP 258  
 DB 459 -RASVDTGSHPASAI-----YSVETINDGNFHVILLTLDSSLSLSVDGG 502  
 QY 259 VPDGYSHASQORVLSFN-----LGDMPRGKVKVTFVEFCFQRRGQITNVAITVTCYGGHKCS 314  
 DB 503 SPKIIITNLKQSTLNFDSPLYVGGM-PGKNNVASLRQAPQNG-----TSFHGCI 551  
 QY 315 ANVTTVVNEPCVQVNISGADWSYVCKPVEYSISVSNPGDLVLDHVVQIDTLPSPGVTVLEA 374

Db 552 RN-----LYINSELQDFRKV--PMQTGI-----L 573  
Qy 375 PGEICCNKVVWRIKEMCPGETLQFLKLVVKAQVPGRTNQAVVTSENCGTCTSCAETTT 434  
Db 574 PCEPCHRRK-----CAHGTQ-----PSSQSGFTCECEE--- 603  
Qy 435 HWKGLAATHMVCVLDNDPICVGVNTVYRICV 465  
Db 604 GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 629

## RESULT 6

US-09-025-596-41  
; Sequence 41, Application US/09025596  
; Patent No. 6340463  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, William M.  
; APPLICANT: Stratton, Charles W.  
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE  
; FILE REFERENCE: VDB98-01  
; CURRENT APPLICATION NUMBER: US/09/025,596  
; EARLIER FILING DATE: 1998-02-18  
; EARLIER APPLICATION NUMBER: 08/911,593  
; EARLIER FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 60/023,921  
; EARLIER FILING DATE: 1996-08-14  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 41  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-025-596-41

Query Match 4.0%; Score 119; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.5e-05;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 418 TSESNCGTCTSCAETTHWK 437  
Db 1 TSESNCGTCTSCAETTHWK 20

## RESULT 7

US-09-191-647-2  
; Sequence 2, Application US/09191647  
; Patent No. 6046015  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; TITLE OF INVENTION: Tessler-Lavigne, Marc  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/191,647  
; EARLIER FILING DATE: 1998-11-13  
; EARLIER APPLICATION NUMBER: 60/065,544  
; EARLIER FILING DATE: 1997-11-14  
; EARLIER APPLICATION NUMBER: 60/081,057  
; EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: human  
US-09-191-647-2

Query Match 4.0%; Score 117.5; DB 3; Length 1525;

Best Local Similarity 20.3%; Pred. No. 0.074;  
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;  
Qy 72 GAFCDKEFYCEEGRCQ-----PVEAQQESC---YGRLY-----SYKVND 109  
Db 1066 GEHCIDIDFDCQDNKCKNGAHTDAVNGYTCICPEGYSGLFCFEPMPVLPRTSPCDNFD 1125  
Qy 110 C-----NVEICQSVPEYATVGSPPYIEILAIGRKDCVDV-----ITOQLP 150  
Db 1126 CONGAQCIIVRINEPICQCLPGYQGEKCLVSVNFINKESYLQIPSAKVRPQTNITLQI- 1184  
Qy 151 CEAEFVSSDPETPTSDGKLWKKIDRLGAGDCKIKVWVKPLKEGCCFPAATVCAPEL- 209  
Db 1185 -----ATDEDSGILLYK-----GDKDHIAV-----ELY 1207  
Qy 210 -----RSYTKGGOAICIKQEGPDCACLRCPVCYKTEVNTGS-----AI--ARNVTV 255  
Db 1208 RGRVRASYDTGSHPSAI-----YSYETINDGNFHVIELLALDQSLSLV 1252  
Qy 256 DNPVPDGYSHASGORVLSFN-----LGDMPRGKKVFTVEFCPPRRGQITNVATVTCGGH 311  
Db 1253 DGGNPKIITNLKQSTLNFDSPLYVGGM-PGKSNVASLRAQPGQNG-----TSFH 1301  
Qy 312 KCSANVTTVVNEPCVOVNISSADMSYCKPVEYSISVNPGLVLJHDVVIQDTLPSTVTV 371  
Db 1302 GCIRN-----LYINSELQDFQV--PMQTGI----- 1325  
Qy 372 LEAPGGEICCNKVVWRIKEMCPGETLQFLKLVVKAQVPGRTNQAVVTSENCGTCTSCAE 431  
Db 1326 --LPGCEPCHKKV-----CAHGTQ-----PSSQAGFTCEQOE 1356  
Qy 432 TTHWKGAAATHMVCVLDNDPICVGVNTVYRICV 465  
Db 1357 --OWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382

## RESULT 8

US-09-540-245A-2  
; Sequence 2, Application US/09540245A  
; Patent No. 6270984  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessler-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,245A  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: human  
US-09-540-245A-2

Query Match 4.0%; Score 117.5; DB 4; Length 1525;  
Best Local Similarity 20.3%; Pred. No. 0.074;  
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;

Qy 72 GAFCDKEFYCEEGRCQ-----PVEAQQESC---YGRLY-----SYKVND 109  
Db 1066 GEHCIDIDFDCQDNKCKNGAHTDAVNGYTCICPEGYSGLFCFEPMPVLPRTSPCDNFD 1125  
Qy 110 C-----NVEICQSVPEYATVGSPPYIEILAIGRKDCVDV-----ITOQLP 150  
Db 1126 CONGAQCIIVRINEPICQCLPGYQGEKCLVSVNFINKESYLQIPSAKVRPQTNITLQI- 1184

Db 1253 DGNPKIITNLSKOSTLNFDSPLYVGM-PGKSNVSLRQAPGONG-TSEH 1301  
QY 312 KCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVNSPGDLVLHVDVVIQDTLPSCVTV 371  
Db 1302 GCIRN-----LYINSELQDFQKV--PMOTGI-----1325  
QY 372 LEAPGEICCNKVVWRKEMCPGETLQFLKLVKAQVPGREFTNOVAVTSESCGTCTSCAE 431  
Db 1326 --LPGCEPCHKV-----CAHGTQ-----PSSOAGFTCECQE 1356  
QY 432 TTHWKGLAATHCMCVLDTNDPICVGTNTYRICV 465  
Db 1357 ---GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382

RESULT 10  
PCT-US95-11684-4  
; Sequence 4, Application PC/TUS9511684  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE  
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING  
; NUMBER OF INVENTION: AND USING SAME  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; STREET: 10666 North Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: PCT/US95/11684  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/308,359  
; FILING DATE: 16-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April C. 33,950  
; REGISTRATION NUMBER: BEC0019P  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1810 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-11684-4

Query Match 3.9%; Score 115.5; DB 5; Length 1810;  
Best Local Similarity 21.1%; Pred. No. 0.15;  
Matches 127; Conservative 65; Mismatches 204; Indels 205; Gaps 37;

QY 82 CEEGRCPQVEAQOESCGRLYSVKNVNDNVEIC-----OSVPEYATVCSPPYIEL 133  
Db 357 CENGLC-----VCEHGF-----VGDDCSQKRCRPTCNRRGRCVDRGCVCHEGY-----399  
QY 134 AIGKDKCDV-----VITQQLPCEAEFVSSDPETPTSDGKLVMKIDRLGAGDK 182  
Db 400 -LG-EDGELRCPNDCHNRGRCINGOCVCEGFIGED-----CGELRCPND 443

QY 151 CEAEFVSSDPETPTSDGKLVMKIDRLGAGDKCKITVWVKPLKRGCCFTAAATVACAPDEL- 209  
Db 1185 -----ATDEDSGILLYK-----GKDHIAV-----ELY 1207  
QY 210 -----RSYTKGQPAICIKQEGPCACILRCPVCYKIEVNTGS-----AI--ARNVTV 255  
Db 1208 RGRVRSYDTGSHPASAI-----YSVETINDGNFHIIVELLALDQSLSLV 1252  
QY 256 DNPVDPGYSHASGORVLSFN-----LGDMPRGDKKVTVEFCPPORRGQITNATVYCCGGH 311  
Db 1253 DGNPKIITNLSKOSTLNFDSPLYVGM-PGKSNVSLRQAPGONG-----TSEH 1301  
QY 312 KCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVNSPGDLVLHVDVVIQDTLPSCVTV 371  
Db 1302 GCIRN-----LYINSELQDFQKV--PMOTGI-----1325  
QY 372 LEAPGEICCNKVVWRKEMCPGETLQFLKLVKAQVPGREFTNOVAVTSESCGTCTSCAE 431  
Db 1326 --LPGCEPCHKV-----CAHGTQ-----PSSOAGFTCECQE 1356  
QY 432 TTHWKGLAATHCMCVLDTNDPICVGTNTYRICV 465  
Db 1357 ---GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382

RESULT 9  
US-09-540-153-2  
; Sequence 2, Application US/09540153  
; Patent No. 6270995  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,153  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/191,647  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: human  
US-09-540-153-2

Query Match 4.0%; Score 117.5; DB 4; Length 1525;  
Best Local Similarity 20.3%; Pred. No. 0.074; Indels 197; Gaps 23;  
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;

QY 72 GAFCDREFYCEGRQ-----PVEAQOESC-----YGRLY-----SVKVNDD 109  
Db 1066 GEHCIDIDFDDCCQNKANGAHCCTDAVNGYTCICPEGYSGLFCEFPMPVLPRTSPCDNFD 1125  
QY 110 C-----NVEICQSVPEYATVGPYPPIELAIATGKDCVDV-----ITQQLP 150  
Db 1126 CQGAOCIVRINEPICQLPGYQGEKCEKLVSVNFINKESYLQIPSAKVRPQNTILQI- 1184  
QY 151 CEAEFVSSDPETPTSDGKLVMKIDRLGAGDKCKITVWVKPLKRGCCFTAAATVACAPEL- 209  
Db 1185 -----ATDEDSGILLYK-----GKDHIAV-----ELY 1207  
QY 210 -----RSYTKGQPAICIKQEGPCACILRCPVCYKIEVNTGS-----AI--ARNVTV 255  
Db 1208 RGRVRSYDTGSHPASAI-----YSVETINDGNFHIIVELLALDQSLSLV 1252  
QY 256 DNPVDPGYSHASGORVLSFN-----LGDMPRGDKKVTVEFCPPORRGQITNATVYCCGGH 311

```
QY 183 KKITVWVKPKLKGCCFTAAATVC-----ACPELRSYTKCQPAICIKQE-----GP 227
Db 444 CQ-----QGRGICNGQCEHGFEDGELURCNDNCNHSGRVCYNGQCVCDGEGTGE 495
QY 228 DCACLRCP-VCYKTEVNTGSAIARNVTVDNVPDGY-----SHASGORVLS 273
Db 496 DCGELRCPNDCH-----NRGRVBCRCVCDN-----GFMGEDCGELSCPNCHQHGRVCV-- 544
QY 274 FNLGDMRPGDKKVTVEFCPPQRQIINVTATVYCGHKKSANVTTVVNEPCV-QVNISG 332
Db 545 ----DGRVCHEGFTGEDCRERS-----C-PNDCN-NVGRVCEGVCCEGYMG 587
QY 333 ADNSYCKPVEYSISVSNPGLDVL-----HDVVI-----QDTLP---S 367
Db 588 IDCSDVSPPT--GLTNTVNTDKTNLEWHENLVNEVLYTVYVTSGLGLDLQFTVPGNQ 645
QY 368 GVTVLE-APGGE-----ICCNK-----VWRIKEMCPG-ETIQFKLVKAQVPGR-- 410
Db 646 SATIHELEPGVEFIRVAILKNNKIPVSARVATYLPAPGLKFKRSVRETSVQVEWDPL 705
QY 411 -----FTNOVAVTSESNGCTCTSCAETTHWKGLA-----ATHMCVLDNDPI 453
Db 706 SISFDGWLTVFRNMOKDDNGDITSLKRPETSYMQPGLAPGOQYVNSLHIVKNNTRGP- 764
QY 454 CVGENTVY--RICVNRGSAED-TNVSLLKFSKELOPIASSGPTKG--TISGNTVYVD- 507
Db 765 --GLSRVITTKLDAPSQTEAKDVTDTALITWSKPLAEIEGIELTYGPKDVPDGRDITDL 822
QY 508 -----ALPKLGSKESVEFSVTL---KGIAPGDARGAAILSSDTLTSP-----VSDTEN 552
Db 823 SEDENOYSIGNLRPHTEVEYEVTLISRRGDMESDPAKEVFVTD--LDAPRNLKRVSDTN 880
QY 553 T 553
Db 881 S 881

RESULT 11
US-08-404-665-4
; Sequence 4, Application US/08404665
; Patent No. 5591583
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,665
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugic, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3341
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1358 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-404-665-4
```

```
Query Match 3.88; Score 112.5; DB 1; Length 1358;
Best Local Similarity 19.28; Pred. No. 0.18;
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;
```

```
QY 23 SGGTEAAVAESLITKIIVASAEKPPAPVMTAKKVLVRRNKQPVQEKSRGAFCDKEFPYC 82
Db 80 SSGLE-----ASAQE-----VSAEDETALAEYMGQTSDBESQVTFTHRINFPK 122
QY 83 EGRGCPVEAQOESCGRLYSVK-----VNDNCNVEICQSVPEYATVGYSPYIILAIGK 137
Db 123 KACPCESSAQVLOELLSEIEMLEREYSLVLRDQCNAACQ-----ESAATGQ 168
QY 138 KDCVDVVIITQOLPCEAEFVSSDPETPTSDGKLVMKIDRLGA-----GDKCKITVWVK 190
Db 169 LDVI-----PHCSGHGNFSEFCGCGICNEGWFGKNCSEPYC-- 204
QY 191 PL---KEGCGFTAATVC-----ACPELRSYTKCQPAICIKQE-----GPDCACL 232
Db 205 PLOCSRGVGVGQCCICDSEYSGDDCSELKCPDTCSSRGLCVGDCVCEBPYTGDCREL 264
QY 233 RCP-VCYKIEVYNTGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVTVEF 291
Db 265 RCPGDCSGKRCATGTCLCE-----EGY-----VGED 291
QY 292 CPOR-----RGQITNVATVYCG--GHKCSANVTTVVNEPCVOVNISG-----AD 334
Db 292 CGORCLNACSGRGOCEGLCVCEGYQGPCSAYA-----PPEDLRVAGISDRSIELE 345
QY 335 WSYCKPVEYSISVSNPGDLVLDVVIODTLP---SGVTLEAPG---EICCCKVWVRI 388
Db 346 WDGPMATVEYVSIQ---PTALGGLQLOQRPDOWSGVTTITELEPLTYNISVYAVISNI 402
QY 389 -----KEMCPGETLOFKLV-----VKAQV 407
Db 403 LSLPITAKVATHLSPTQGLQFKTITETTVQVOWEPFSEFDCGWEISFIPKNEGVIQAV 462
QY 408 PGRET-----NOVAVTSESNGCTCTSCAETTHWKGLAATHMCVLDNDP 452
Db 463 PSDVTSFNQTLGKPEEYIVNVVALKEQARSPTTSASVSTVID-----GPTOILVRDVS 518
QY 453 ICVGENTVYRICVNRGSAEDTNVSLIKF-----SKELQPIASSGPTKGTISGN 502
Db 519 VAFVE-----WIPPRAKVD-----FILLKYGLVGGEGGRTFRLQPLPSQYSVALRPGS 568
QY 503 TVVFDALPKLGSKESVEFSVTLKGIAPGDARGAAILSSDTLTSPVSDTEN 553
Db 569 RYEVSVSAVRGTNES--DSATTQFTTIDAPKLNLRVGSRTATSLDLEWDNS 617
```

```
RESULT 12
US-08-404-671-4
; Sequence 4, Application US/08404671
; Patent No. 5635360
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
```

ZIP: 07417  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/404.671  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-3341  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1358 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-404-671-4

Query Match 3.8%; Score 112.5; DB 1; Length 1358;  
Best Local Similarity 19.2%; Pred. No. 0.18;  
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;

Qy 23 SGGIEAAVAESLTKIVASAETKPAVPMTAKKVLRRNKQPVQKSRGAFCDKEFYPC 82  
Db 80 SSGLE-----ASAQEQ-----VSAEDETLAEYMGQTSDEHSQVTFTHRINFPK 122  
Qy 83 EEGRCQPVQAQESYGRLYSVK-----VNDNCNVEICQSVPEYATVGSYPPIELAIK 137  
Db 123 KACPSSSAQVQLLESLRIEMLEREVSVLRQCNANCCQ-----ESAATGQ 168  
Qy 138 KCDVVDVITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA-----GDKCKITVWVK 190  
Db 169 LDYI-----PHCSGHGNEFSESCGICNEGWFKNKSEPYC-- 204  
Qy 191 PL---REGCFTAAVTC-----ACPELRSYTRKCGQPAICIKOE-----GPDCACL 232  
Db 205 PLGCSRGVCDGQCICDSEYSGDDCSELRCPTDCSSRGLCVGCEVCEPYTGDCREL 264  
Qy 233 RCP-VCYKIEVNTGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVEF 291  
Db 265 RCPGDCSGKRCATGTCICE-----EGY-----VGED 291  
Qy 292 CPQR-----RGQITNVAITYCG--GHKCSANVTTVNNEPCVQVNISSG-----AD 334  
Db 292 CGORQCLNACSGRGQCEGLCVCEGYGQPCSAVA-----PPEDLRVAGISDRSIELE 345  
Qy 335 WSYVCKPVEYSISVSNPGLVLDHVVIOQTLF---SGVTLEAPGG---EICCNKVVWRI 388  
Db 346 WDGPMAVTEYVISYQ---PTALGGQLQORVPGDWSGVITTELEPGLTYNISVYAVISNI 402  
Qy 389 -----KEMCPGETLOFLV-----VKAQV 407  
Db 403 LSLPITAKVATHLSLTPGQLOFRTITETTVQVQWEPFSEFSDGWEISIPKNNEGGVIAQV 462  
Qy 408 PGRFT-----NOVAVTSESCGTCTSCAETTHWKGLAATHMCVLDTNDP 452  
Db 463 PSDVTSFNQTLKPGBEYITNVVVALKEQARSPTTSASVSTVID---GPTQILVRVSDT 518  
Qy 453 ICVENTVYRICVTNRGSADNTVSLIKF-----SKELOPTASSSGPTKGTISGN 502  
Db 519 VAFVE-----WIPPRKVD---FILLKYLGVGEGGRTTFRLOPLSQYSVQALRPS 568  
Qy 503 TVVFDALPKLGSKEVSFVTKGIAPDAGARGEAILSDSLTSPVSSTENT 553  
Db 569 RYEVSVSAVRGTNES--DSATQFTTEIDAPKNLRVGSRTATSLDLEWDS 617

RESULT 13  
US-08-404-781-4  
Sequence 4, Application US/08404781  
Patent No. 5681931  
GENERAL INFORMATION:  
APPLICANT: Reid, Robert A.  
APPLICANT: Ackley, Rhonda L.  
APPLICANT: Hemperly, John J.  
TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and  
STREET: 1 Becton Drive  
CITY: Franklin Lakes  
STATE: NJ  
COUNTRY: US  
ZIP: 07417  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/404.781  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fugit, Donna R.  
REGISTRATION NUMBER: 32,135  
REFERENCE/DOCKET NUMBER: P-3341  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1358 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-404-781-4

Query Match 3.8%; Score 112.5; DB 1; Length 1358;  
Best Local Similarity 19.2%; Pred. No. 0.18;  
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;

Qy 23 SGGIEAAVAESLTKIVASAETKPAVPMTAKKVLRRNKQPVQKSRGAFCDKEFYPC 82  
Db 80 SSGLE-----ASAQEQ-----VSAEDETLAEYMGQTSDEHSQVTFTHRINFPK 122  
Qy 83 EEGRCQPVQAQESYGRLYSVK-----VNDNCNVEICQSVPEYATVGSYPPIELAIK 137  
Db 123 KACPSSSAQVQLLESLRIEMLEREVSVLRQCNANCCQ-----ESAATGQ 168  
Qy 138 KCDVVDVITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA-----GDKCKITVWVK 190  
Db 169 LDYI-----PHCSGHGNEFSESCGICNEGWFKNKSEPYC-- 204  
Qy 191 PL---REGCFTAAVTC-----ACPELRSYTRKCGQPAICIKOE-----GPDCACL 232  
Db 205 PLGCSRGVCDGQCICDSEYSGDDCSELRCPTDCSSRGLCVGCEVCEPYTGDCREL 264  
Qy 233 RCP-VCYKIEVNTGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVEF 291  
Db 265 RCPGDCSGKRCATGTCICE-----EGY-----VGED 291  
Qy 292 CPQR-----RGQITNVAITYCG--GHKCSANVTTVNNEPCVQVNISSG-----AD 334  
Db 292 CGORQCLNACSGRGQCEGLCVCEGYGQPCSAVA-----PPEDLRVAGISDRSIELE 345  
Qy 335 WSYVCKPVEYSISVSNPGLVLDHVVIOQTLF---SGVTLEAPGG---EICCNKVVWRI 388  
Db 346 WDGPMAVTEYVISYQ---PTALGGQLQORVPGDWSGVITTELEPGLTYNISVYAVISNI 402

```
QY 389 -----KMCPEGTLQFKLV-----VKAQV 407
      : : : : :
Db 403 LSLPITAKVATHLSTPOGLQFKTITETTVQVQEPFSEFSGWEISFIPKNNEGGVAQ 462
      : : : : :
QY 408 PCRFT-----NOVAVTSNCGTCTSCAETTHWKGLAATHMCAVLDTNDP 452
      : : : : :
Db 463 PSDVTSFNGTLKPGEEYIVNVALKQEARSPPTSASVSVID---GPTQILVRVSDT 518
      : : : : :
QY 453 ICVGENTYVYRICVTHRGSAEDTNVSLIKF-----SKELOPIASSGPTKGIISGN 502
      : : : : :
Db 519 VAFVE-----WIPRAKVD-----FILLKYLGVGEGGRTRFRLQPLPSYSVQALRPGS 568
      : : : : :
QY 503 TVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSDTLTSPVSDTENT 553
      : : : : :
Db 569 RYEVSVSAVRGTNES--DSATTQFTTEIDAPKNLRVGSRTATSLDLEWDS 617

RESULT 14
US-08-682-517-15
: Sequence 15, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1222 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-15

Query Match 3.8%; Score 111.5; DB 2; Length 1222;
Best Local Similarity 20.6%; Pred. No. 0.19;
Matches 114; Conservative 65; Mismatches 209; Indels 165; Gaps 24;

QY 99 GRLYSVKVNDDCNVEICQSVPEYATVGSYPPIEILAIKDKDCVDVWITQOLPCEAEFVSS 158
      : : : : :
Db 523 GREYKIVKD-----KAGNLAKNEIVNVAFNEDKDRVIS--TVTNKAFVDT 566
      : : : : :
QY 159 DPET-----TPTSD-GKLWVKIDRLGAGDKCKITVWV----KPLKEGCGCFTA 200
      : : : : :
Db 567 DPDTAVYFTGDKAKOISVKTNKDGATEVIGSDTVNDYATPIAWIDINTSDAKQGLDEG 626
      : : : : :
QY 201 ATVCACP-----ELRSYTKCGOPAIKIQEGPCACLRCPVCYKIEVNV---244
      : : : : :
Db 627 EPKAVAPISYFOAPYLDGSAIKAYKKSDLNKAFTKFDGSETA-----VFAAELVNSGK 680
      : : : : :
QY 245 --TGSATARNVVDNVPDPGVSHASQORVLSFNLDGMRPDKKYFTVEFCPOBRRGOITNV 302
      : : : : :
Db 681 KVTGTSIKK-----ATYTIYNTGANDIKVDNQVISPNSRYTVTYE 720
      : : : : :
QY 303 ATVTYCGGHKCSA---NVTVV-----VNEPCVQVNISGADWSYVCKPVEYSISVSNP 351
      : : : : :
Db 721 ATLSSTGTVITPAKMLEVTSVDGKTTAVKIATGIAVNTDGKDYAFTAKAATATFTATNE 780
      : : : : :
QY 352 GDLVLHDVVIQ-DTLPSC-----VTVLEAPGGEICCNKVVWR 387
      : : : : :
Db 781 VPNSYTGATQNTADSGSNSIWFAGKNPKVYAGVSGKTYKYFGANGNEV-FGEAAWE 839
      : : : : :
QY 388 -----IKEMCPGETLQFLVVKVQVPGRTNQAVTS-----ESNCGTCTSC 429
```

```
Db 840 ALLTQYATEGOKVTISYVNDGDTVTFKVISAVNSSTEATKPVAPTTTGAULTLTPA 899
      : : : : :
QY 430 A-----ETTHMKGLAATHMCAVLDTNDPICVGENTVYVYICVTHRGSAEDTNVSLIL--- 480
      : : : : :
Db 900 AGGLVDLTATNTLIGIS---LADADLN---VSATVTDVATVSLKDSA--NNSLSLTLVET 951
      : : : : :
QY 481 ----KESKELOPIASSGPTKGIISGNVTVF--DALPKLGSKEVSFVTLKGIAPGDARG 534
      : : : : :
Db 952 GANTGVFATTVQAGTLSLTAGTL---TVTYADAKNAAGVAENITASVTLK-----KT 1001
      : : : : :
QY 535 EAILSSDTLSPV 547
      : : : : :
Db 1002 TGAITSDTFTQGV 1014

RESULT 15
US-08-682-517-9
: Sequence 9, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1252 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-9

Query Match 3.8%; Score 111.5; DB 2; Length 1252;
Best Local Similarity 20.6%; Pred. No. 0.2;
Matches 114; Conservative 65; Mismatches 209; Indels 165; Gaps 24;

QY 99 GRLYSVKVNDDCNVEICQSVPEYATVGSYPPIEILAIKDKDCVDVWITQOLPCEAEFVSS 158
      : : : : :
Db 553 GREYKIVKD-----KAGNLAKNEIVNVAFNEDKDRVIS--TVTNKAFVDT 596
      : : : : :
QY 159 DPET-----TPTSD-GKLWVKIDRLGAGDKCKITVWV----KPLKEGCGCFTA 200
      : : : : :
Db 587 DPDTAVYFTGDKAKOISVKTNKDGATEVIGSDTVNDYATPIAWIDINTSDAKQGLDEG 656
      : : : : :
QY 201 ATVCACP-----ELRSYTKCGOPAIKIQEGPCACLRCPVCYKIEVNV---244
      : : : : :
Db 657 EPKAVAPISYFOAPYLDGSAIKAYKKSDLNKAFTKFDGSETA-----VFAAELVNSGK 710
      : : : : :
QY 245 --TGSATARNVVDNVPDPGVSHASQORVLSFNLDGMRPDKKYFTVEFCPOBRRGOITNV 302
      : : : : :
Db 711 KVTGTSIKK-----ATYTIYNTGANDIKVDNQVISPNSRYTVTYE 750
      : : : : :
QY 303 ATVTYCGGHKCSA---NVTVV-----VNEPCVQVNISGADWSYVCKPVEYSISVSNP 351
      : : : : :
Db 751 ATLSSTGTVITPAKMLEVTSVDGKTTAVKIATGIAVNTDGKDYAFTAKAATATFTATNE 810
      : : : : :
QY 352 GDLVLHDVVIQ-DTLPSC-----VTVLEAPGGEICCNKVVWR 387
      : : : : :
Db 811 VPNSYTGATQNTADSGSNSIWFAGKNPKVYAGVSGKTYKYFGANGNEV-FGEAAWE 869
      : : : : :
QY 388 -----IKEMCPGETLQFLVVKVQVPGRTNQAVTS-----ESNCGTCTSC 429
      : : : : :
Db 870 ALLTQYATEGOKVTISYVNDGDTVTFKVISAVNSSTEATKPVAPTTTGAULTLTPA 929
```

|    |      |  |   |      |
|----|------|--|---|------|
| QY | 430  | A  | -----ETTHWKGGLAATHCMVCLTDNPICVGTGNTYVYICVTNKGSRADTNVSLIL--- | 480  |
| DB | 930  | AGCLVDLTATNTLGIS---  | LADALN-----VSAATVDTATVSLKDSA--NNLSLSLTJET                   | 981  |
| QY | 481  | -----KFSKEQLPIASGPFKCTIGNTWVF--DALPKLSKGSVEFSTVLKGIAPGDARG | 534   |      |
| DB | 982  | GANTGVATTVQAGTGLUSSLTAGTL---                               | TVTYADAKNAAGVAENITASVTLK-----KT                             | 1031 |
| QY | 535  | EALLSDDTLSPV   | 547   |      |
| DB | 1032 | TGAIISDTFTQGV  | 1044  |      |

Search completed: May 25, 2002, 22:15:31  
Job time: 5605 sec





Tue May 28 08:55:19 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 20:48:46 ; Search time 56.2 seconds  
(without alignments)  
950.635 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 2951  
Sequence: 1 MSKLRRVVVLTALTSASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description          |
|------------|--------|-------------|--------|----------|----------------------|
| 1          | 2951   | 100.0       | 556    | 1 S12602 | 60K cysteine-rich    |
| 2          | 2951   | 100.0       | 556    | 2 A86560 | 60 kDa Cysteine-rich |
| 3          | 2553.5 | 86.5        | 557    | 1 B39439 | 60K cysteine-rich    |
| 4          | 2488   | 84.3        | 558    | 2 J5204  | 60K cysteine-rich    |
| 5          | 2176   | 73.7        | 554    | 2 C81671 | 60 kDa outer membr   |
| 6          | 2162.5 | 73.3        | 547    | 2 B43584 | 60K cysteine-rich    |
| 7          | 2160.5 | 73.2        | 553    | 2 D71515 | 60K cysteine-rich    |
| 8          | 2129.5 | 72.2        | 547    | 1 A32244 | 60K cysteine-rich    |
| 9          | 151.5  | 5.1         | 1474   | 2 F69009 | probable membrane    |
| 10         | 135    | 4.6         | 2233   | 2 T28669 | surface protein 51   |
| 11         | 130.5  | 4.4         | 756    | 2 D75622 | hypothetical prote   |
| 12         | 130.5  | 4.4         | 1408   | 2 S16148 | gene serrate prote   |
| 13         | 129    | 4.4         | 772    | 2 A48822 | protein-glutamine    |
| 14         | 127.5  | 4.3         | 998    | 2 T25489 | conserved hypothet   |
| 15         | 127    | 4.3         | 693    | 2 T25878 | hypothetical prote   |
| 16         | 124    | 4.2         | 2195   | 2 T34264 | hypothetical prote   |
| 17         | 123.5  | 4.2         | 2139   | 2 A35672 | crumbs protein - f   |
| 18         | 123    | 4.2         | 647    | 2 A43902 | tenascin - eastern   |
| 19         | 123    | 4.2         | 1295   | 2 A32501 | glp1 protein precu   |
| 20         | 122    | 4.1         | 491    | 2 B75554 | hypothetical prote   |
| 21         | 122    | 4.1         | 2910   | 2 T42214 | otogelin - mouse     |
| 22         | 121    | 4.1         | 641    | 2 D71170 | hypothetical prote   |
| 23         | 120.5  | 4.1         | 354    | 2 T22274 | hypothetical prote   |
| 24         | 120    | 4.1         | 1356   | 2 A45445 | Janusin precursor,   |
| 25         | 120    | 4.1         | 2533   | 2 T28675 | alpha-51D immobili   |
| 26         | 120    | 4.1         | 2533   | 2 T28674 | hypothetical prote   |
| 27         | 119    | 4.0         | 1599   | 2 T16210 | probable peptidogl   |
| 28         | 118    | 4.0         | 1612   | 2 AB1347 | restrictin precurs   |
| 29         | 117.5  | 4.0         | 1353   | 1 JH0675 |                      |

|    |       |     |      |          |                     |
|----|-------|-----|------|----------|---------------------|
| 30 | 117.5 | 4.0 | 1658 | 2 D75489 | hypothetical prote  |
| 31 | 116.5 | 3.9 | 565  | 2 T16408 | hypothetical prote  |
| 32 | 115.5 | 3.9 | 577  | 2 A60501 | thrombomodulin pre  |
| 33 | 115.5 | 3.9 | 1810 | 1 A32230 | tenascin precursor  |
| 34 | 115   | 3.9 | 2543 | 2 T31687 | surface antigen - p |
| 35 | 114.5 | 3.9 | 2918 | 2 A54105 | fibrillin-2 precur  |
| 36 | 113.5 | 3.8 | 626  | 2 T27319 | hypothetical prote  |
| 37 | 113.5 | 3.8 | 2946 | 2 T15840 | hypothetical prote  |
| 38 | 113   | 3.8 | 2907 | 2 A57278 | secreted leucine-r  |
| 39 | 112.5 | 3.8 | 1025 | 2 T4626  | hypothetical prote  |
| 40 | 112.5 | 3.8 | 1620 | 2 T27283 | hypothetical prote  |
| 41 | 112.5 | 3.8 | 1806 | 2 T23298 | hypothetical prote  |
| 42 | 112.5 | 3.8 | 2271 | 2 F90073 | hypothetical prote  |
| 43 | 112   | 3.8 | 2471 | 2 A49128 | cell-fate determin  |
| 44 | 112   | 3.8 | 2871 | 2 A55624 | fibrillin-1 precur  |
| 45 | 111.5 | 3.8 | 1746 | 1 S19694 | tenascin precursor  |

ALIGNMENTS

RESULT 1

S12602

60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydoiphila pneum  
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae  
C:Date: 30-Sep-1993 #sequence-revision 27-Jun-1994 #text\_change 11-May-2000  
C:Accession: S12602; H72063; A81604  
R:Watson, M.W.; Al-Mahdawi, S.; Lamden, P.R.; Clarke, I.N.  
Nucleic Acids Res. 18, 5299, 1990  
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of  
A:Reference number: S12602; MUID:90384850  
A:Accession: S12602  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-556 <WAT>  
A:Cross-references: GB:X53511; NID:g550564; PIDN:CAA37590.1; PID:g550566  
A:Experimental source: isolate IOL-207  
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: H72063  
A:Molecule type: DNA  
A:Residues: 1-556 <ARN>  
A:Cross-references: GB:AE001640; GB:AE001363; NID:g4376845; PIDN:AAD18697.1; PID:g437  
A:Experimental source: strain CML029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: A81604  
A:Molecule type: DNA  
A:Residues: 1-556 <REA>  
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Comment: This protein is associated with the differentiation of reticulate bodies  
. It may also be an important virulence factor.  
C:Genetics:  
A:Gene: omcB; CP0195  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <MAT>

Query Match 100.0%; Score 2951; DB 1; Length 556;  
Best Local Similarity 100.0%; Pred. No. 2.9e-197;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 MSKLRRVVVLTALTSASCASGIEAAVSLTKIVASAEKPAVPMTAKKRLVLR 60  
|||||

Db 1 MSLKLRVVTVLALTSMA SCFASGIEAAVAESLTKIVASAETKPAVPMTAKKVLVR 60  
QY 61 RNKQPVQKSRGAFCDKEFYPCCEGRCPVQEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
Db 61 RNKQPVQKSRGAFCDKEFYPCCEGRCPVQEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
QY 121 YATVGSPPYPIEIIAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
Db 121 YATVGSPPYPIEIIAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 241 EVVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
QY 301 NVATVTCGGHKCSANVTTVVNEPCQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
Db 301 NVATVTCGGHKCSANVTTVVNEPCQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
QY 361 IDOTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
Db 361 IDOTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
QY 421 SNGCTCTCAETTHWKGLAATHMCVLDNDPICVGVNTVYRICVTNNGSAEDTNVSLIL 480  
Db 421 SNGCTCTCAETTHWKGLAATHMCVLDNDPICVGVNTVYRICVTNNGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEVFSVTLKGAPGDARGEAILSS 540  
Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEVFSVTLKGAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHVV 556  
Db 541 DTLTSPVSDTENTHVV 556

RESULT 2  
60 kDa Cysteine-rich omp [imported] - Chlamydoiphila pneumoniae (strain J138)  
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
A:Accession: A86560  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Ouchi, K.; Shiba, T.; Is  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: A86560  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-556 <STO>  
A:Cross-references: GB:BA000008; NID:g8978928; PIDN:BAA98763.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: OmcB  
C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 100.0%; Score 2951; DB 2; Length 556;  
Best Local Similarity 100.0%; Pred. No. 2.9e-197;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLKLRVVTVLALTSMA SCFASGIEAAVAESLTKIVASAETKPAVPMTAKKVLVR 60  
Db 1 MSLKLRVVTVLALTSMA SCFASGIEAAVAESLTKIVASAETKPAVPMTAKKVLVR 60  
QY 61 RNKQPVQKSRGAFCDKEFYPCCEGRCPVQEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
Db 61 RNKQPVQKSRGAFCDKEFYPCCEGRCPVQEAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
QY 121 YATVGSPPYPIEIIAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180

Db 121 YATVGSPPYPIEIIAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180  
QY 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
Db 181 DKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240  
QY 241 EVVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 241 EVVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
QY 301 NVATVTCGGHKCSANVTTVVNEPCQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
Db 301 NVATVTCGGHKCSANVTTVVNEPCQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360  
QY 361 IDOTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
Db 361 IDOTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSE 420  
QY 421 SNGCTCTCAETTHWKGLAATHMCVLDNDPICVGVNTVYRICVTNNGSAEDTNVSLIL 480  
Db 421 SNGCTCTCAETTHWKGLAATHMCVLDNDPICVGVNTVYRICVTNNGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEVFSVTLKGAPGDARGEAILSS 540  
Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEVFSVTLKGAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHVV 556  
Db 541 DTLTSPVSDTENTHVV 556

RESULT 3  
B39439  
60K cysteine-rich outer membrane protein 1 precursor - Chlamydoiphila psittaci  
C:Species: Chlamydoiphila psittaci, Chlamydia psittaci  
C:Date: 21-Feb-1992 #sequence\_revision 27-Jun-1994 #text\_change 31-Mar-2000  
A:Accession: B39439; S12603  
R:Everett, K.D.E.; Hatch, T.P.  
J. Bacteriol. 173, 3821-3830, 1991  
A:Title: Sequence analysis and lipid modification of the cysteine-rich envelope prote  
A:Reference number: A39439; MUID:91267949  
A:Accession: B39439  
A:Molecule type: DNA  
A:Residues: 1-557 <EVE>  
A:Cross-references: GB:M61116; NID:g144489; PIDN:AAB61619.1; PID:g144491  
A:Experimental source: strain 6BC  
R:Watson, M.W.; Lambden, P.R.; Clarke, I.N.  
Nucleic Acids Res. 18, 5300, 1990  
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of  
A:Reference number: S12603; MUID:90384851  
A:Accession: S12603  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-44, 'A', '46-72', 'E', '74-557' <WAT>  
A:Cross-references: GB:X53512; NID:g40625; PIDN:CAA37592.1; PID:g40627  
A:Experimental source: strain EAE/A32/M  
C:Function:  
A:Description: associated with the differentiation of reticulate bodies into elementa  
o be an important virulence factor.  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status predicted <PRO>  
F:41-557/Product: 60K cysteine-rich outer membrane protein 1 #status predicted <MAT>

Query Match 86.5%; Score 2553.5; DB 1; Length 557;  
Best Local Similarity 85.3%; Pred. No. 1.1e-169;  
Matches 475; Conservative 30; Mismatches 51; Indels 1; Gaps 1;

QY 1 MSLKLRVVTVLALTSMA SCFASGIEAAVAESLTKIVASAETKPAVPMTAKKVLVR 59  
Db 1 MSLKLRVVTVLALTSMA SCFASGIEAAVAESLTKIVASAETKPAVPMTAKKVLVR 59

|                       |  |   |                       |
|-----------------------|--|---|-----------------------|
| Db                    | 1  | MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG    | 60                    |
| Qy                    | 60   | RRNQPVQEKSRGAFCDKFEFPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSP    | 119                   |
| Db                    | 61   | RRNQPVQEKSRGAFCDKFEFPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEISQSP    | 120                   |
| Qy                    | 120  | EVATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA | 179                   |
| Db                    | 121  | EVATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGQ | 180                   |
| Qy                    | 180  | GDCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCACLRCPVCYK   | 239                   |
| Db                    | 181  | GECKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCACLRCPVCYK   | 240                   |
| Qy                    | 240  | IEVNTGSAIARNVTVNDNPVDPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPQRGQI    | 299                   |
| Db                    | 241  | IEVNTGSAIARNVTVNDNPVDPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPQRGKV    | 300                   |
| Qy                    | 300  | TNATVTCYCGGHKCSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLHDV  | 359                   |
| Db                    | 301  | TNATVTCYCGGHKCSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLKLYD   | 360                   |
| Qy                    | 360  | VIOTLPSGVTVLRAPEAGCCCNKVMRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTS     | 419                   |
| Db                    | 361  | VIOTLPSGVTVLRAPEAGCCCNKVMRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTK     | 420                   |
| Qy                    | 420  | ESNCGTCTSCAETHHMKGLAATHMCLVDTNDPICVGENTVYRICVTVNRGSAEDTNVSLI  | 479                   |
| Db                    | 421  | NSDCGTCTSCAETHHMKGLAATHMCLVDTNDPICVGENTVYRICVTVNRGSAEDTNVSLI  | 480                   |
| Qy                    | 480  | LKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAIL    | 539                   |
| Db                    | 481  | LKFSKELQPIASSGPTKGTITGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAIL    | 540                   |
| Qy                    | 540  | SDTLSPVSDTENTHVV  | 556                   |
| Db                    | 541  | SDTLTPVPVADTENTHVV  | 557                   |
| RESULT                | 4  |   |                       |
| JC5204                |  |   |                       |
| A:                    | 60K cysteine-rich outer membrane protein 2 precursor - Chlamydomophila psittaci        |   |                       |
| C:                    | Species: Chlamydomophila psittaci, Chlamydia psittaci                                  |   |                       |
| C:                    | Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 31-Mar-2000              |   |                       |
| C:                    | Accession: JC5204  |   |                       |
| A:                    | Reference number: JC5203; MUID:97075924  |   |                       |
| A:                    | Title: Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: Structu |   |                       |
| A:                    | Residues: 1-558 <HSI>  |   |                       |
| A:                    | Cross-references: GB:U41759; NID:g1783376; PIDN:AAB41143.1; PID:g1783382               |   |                       |
| A:                    | Experimental source: strain GPIC   |   |                       |
| C:                    | Genetics:  |   |                       |
| C:                    | Gene: omp2   |   |                       |
| C:                    | Superfamily: 60K cysteine-rich outer membrane protein                                  |   |                       |
| C:                    | Keywords: membrane protein; virulence  |   |                       |
| F:                    | 1-22/Domain: signal sequence #status predicted <SIG>                                   |   |                       |
| F:                    | 23-40/Domain: propeptide #status predicted <PRO>                                       |   |                       |
| F:                    | 41-558/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>     |   |                       |
| Query Match           | 84.38;   | Score 2488;   | DB 2; Length 558;     |
| Best Local Similarity | 82.11;   | Pred. No. 4.1e-165;   |                       |
| Matches               | 458; Conservative  | 44; Mismatches  | 54; Indels 2; Gaps 2; |
| Qy                    | 1  | MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG    | 59                    |
| Db                    | 1  | MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG    | 60                    |
| Qy                    | 60   | RRNQPVQEKSRGAFCDKFEFPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSP    | 118                   |

|                       |   |   |                       |
|-----------------------|---|---|-----------------------|
| Db                    | 61  | RRNQPVQEKSRGAFCDKFEFPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEISQAV    | 120                   |
| Qy                    | 119   | PEYATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLG | 178                   |
| Db                    | 121   | PEYATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLG | 180                   |
| Qy                    | 179   | AGCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCACLRCPVCY    | 238                   |
| Db                    | 181   | QGERCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCACLRCPVCY  | 240                   |
| Qy                    | 239   | KIEVNTGSAIARNVTVNDNPVDPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPQRGQ    | 298                   |
| Db                    | 241   | KIEVNTGSAIARNVTVNDNPVDPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPQRGK    | 300                   |
| Qy                    | 299   | ITNATVTCYCGGHKCSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLHD  | 358                   |
| Db                    | 301   | ITNATVTCYCGGHKCSANVTTVNNEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLKLYD  | 360                   |
| Qy                    | 359   | VVIOTLPSGVTVLRAPEAGCCCNKVMRIKEMCPGETLQFKLVVKAQVPGRTNOVAVT     | 418                   |
| Db                    | 361   | VVIOTLPSGVTVLRAPEAGCCCNKVMRIKEMCPGETLQFKLVVKAQVPGRTNOVAVK     | 420                   |
| Qy                    | 419   | SESNGTCTSCAETHHMKGLAATHMCLVDTNDPICVGENTVYRICVTVNRGSAEDTNVSL   | 478                   |
| Db                    | 421   | TNSDCGTCTSCAETHHMKGLAATHMCLVDTNDPICVGENTVYRICVTVNRGSAEDTNVSL  | 480                   |
| Qy                    | 479   | ILKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAIL   | 538                   |
| Db                    | 481   | ILKFSKELQPIASSGPTKGTITGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAIL   | 540                   |
| Qy                    | 539   | SSDTLSPVSDTENTHVV   | 556                   |
| Db                    | 541   | SSDTLTPVPVADTENTHVV   | 558                   |
| RESULT                | 5   |   |                       |
| C81671                |   |   |                       |
| A:                    | 60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg) |   |                       |
| C:                    | Species: Chlamydia muridarum, Chlamydia trachomatis MoPn                            |   |                       |
| C:                    | Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000           |   |                       |
| C:                    | Accession: C81671   |   |                       |
| A:                    | Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke |   |                       |
| A:                    | Nucleic Acids Res. 28, 1397-1406, 2000  |   |                       |
| A:                    | Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39 |   |                       |
| A:                    | Reference number: A81500; MUID:20150255   |   |                       |
| A:                    | Accession: C81671   |   |                       |
| A:                    | Status: preliminary   |   |                       |
| A:                    | Molecule type: DNA  |   |                       |
| A:                    | Residues: 1-554 <TET>   |   |                       |
| A:                    | Cross-references: GB:AE002341; GB:AE002160; NID:g17190754; PIDN:AAF39537.1; PID:g71 |   |                       |
| A:                    | Experimental source: strain Nigg (MoPn)   |   |                       |
| C:                    | Genetics:   |   |                       |
| C:                    | Gene: TC0727  |   |                       |
| C:                    | Superfamily: 60K cysteine-rich outer membrane protein                               |   |                       |
| Query Match           | 73.78;  | Score 2176;   | DB 2; Length 554;     |
| Best Local Similarity | 72.88;  | Pred. No. 1.8e-143;   |                       |
| Matches               | 405; Conservative   | 74; Mismatches  | 69; Indels 8; Gaps 3; |
| Qy                    | 1   | MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVLR     | 60                    |
| Db                    | 7   | MNKLIRRAVTVFAVTSVASLFAFGVLETSMAESLSTNVLADTK-AKETTSQDKRKAR     | 65                    |
| Qy                    | 61  | RNKQPVQEKSRGAFCDKFEFPCCEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICQSP    | 120                   |
| Db                    | 66  | KNHQ-----NRTSVVRKEVTAVRDTKA--VEPQSCGCKMTYKVNDDNRNVEICQSP      | 118                   |
| Qy                    | 121   | YATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG | 180                   |
| Db                    | 119   | YATVGSPPYPIELAIKGGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGOG | 178                   |



Tue May 28 08:55:19 2002

Q:Function:  
A:Description: associated with differentiation of reticulate bodies into elementary bodies  
A:Note: essential for the structural integrity of the outer envelope of the elementary body  
Q:Superfamily: 60K cysteine-rich outer membrane protein  
Q:Keywords: membrane protein; virulence  
P:1-28/Domain: signal sequence #status predicted <SIG>  
P:29-46/Domain: propeptide #status predicted <PRO>  
P:47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>

Query Match 73.2%; Score 2160.5; DB 2; Length 553;  
Best Local Similarity 71.7%; Pred. No. 2.2e-142; Indels 15; Gaps 4;  
Matches 401; Conservative 75; Mismatches 68;

QY 1 MSKLIRRVTVLALTSMAFCGAEAAVAESLITKIVASAEETKPAVPMTAKKVLVLR 60  
DB 7 MNKLIRRAVTITPAVTSVASLFSAGVLETSMAESLSTNVISLADTKAK--DNTSHKSKKAR 64  
QY 61 RN---KQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOQESCYGRLYSVKYNDDCNVEICQS 117  
DB 65 KHSKKEPVDK-----KEVAPVHESKA--TGPKQDCFCGRMYTVKVNDRNRVEITQA 114  
QY 118 VPEYATVGSYPPIELAILAIGCKDCVDVITQOLPCEAEFVSDDPETPTSDGKLWVKIDRL 177  
DB 115 VPEYATVGSYPPIELIITATGRKDCVDVITQOLPCEAEFVRSDEATPTADGKLWVKIDRL 174  
QY 178 GAGDKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVC 237  
DB 175 GQGEKSKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICVKOEGPENACLRCPVV 234  
QY 238 YKIEVNTVNGSAIARNVTVDNVPDGYSHASQORVLSFNGLDMRPGDKKVTVEFCPPORRG 297  
DB 235 YKINIVNQGTATARNVVENPVDGYAHSSGQVRVLTTLGDMQGEHRTITVEFCPLKRG 294  
QY 298 QITNVATVTCGGHKCSANVTTVVNEPCVOVNI SGADMSYVCKPVEYSISVSNPGDLVLH 357  
DB 295 RATNIAMVSVCGGHKNFASVTTVINEPCVQVSIAGADWSYVCKPVEYSISVSNPGDLVLR 354  
QY 358 DWIQTDLSPGVTVLEAPGGIICCNKVWRKEMCPGETLOFKLVVKAQVGRNTNOVAV 417  
DB 355 DVVVEDTLSPGVTVLEAAGAQISCNKVWTVKELNPGESLQYKVLVRAQTGQGTNNVVV 414  
QY 418 TSENCGCTCAETTHWKGLAATHMCVLDNDPICVGENTVYRICVTRNGSAEDTNVS 477  
DB 415 KSCSDCGTCTCAETTHWKGVAATHMCVDTCDPCVCGENTVYRICVTRNGSAEDTNVS 474  
QY 478 LILKFSKELQIASSGPTKGTISGNVTVFDALPKLGSKESVEFSTLKGIAPDGARGAI 537  
DB 475 LMLKFSKELQPVSFSGPTKGTITGNTVVVFDLSPLRGLSKRETVEFSTLKAVSAGDARGAI 534  
QY 538 LSSDTLTSFVSDTENTHYV 556  
DB 535 LSSDTLTVFVSDTENTHIY 553

RESULT 8  
A32244  
60K cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia tra  
C:Species: Chlamydia trachomatis  
C:Date: 12-Oct-1989 #sequence\_revision 27-Jun-1994 #text\_change 16-Jul-1999  
C:Accession: A32244; A43584; A36043; A30472; JT0419; S18981; S24277  
J:Allen, J.E.; Stephens, R.S.  
J: Bacteriol. 171, 285-291, 1989  
A:Title: Identification by sequence analysis of two-site posttranslational processing of  
A:Reference number: A32244; MUID:89123030  
A:Accession: A32244  
A:Molecule type: DNA  
A:Residues: 1-547 <ALL>  
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553  
A:Experimental source: strain L2/434/Bu  
A:Note: parts of this sequence, including the amino ends of the precursor and mature pro  
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.  
Infect. Immun. 59, 1196-1201, 1991

A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal  
A:Reference number: A43584; MUID:91147205  
A:Accession: A43584  
A:Molecule type: DNA  
A:Residues: 1-547 <DEL>  
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553  
A:Experimental source: serovar 2, strain L2/434/Bu  
A:Wahlberg, J.; Lundberg, J.; Hultman, T.; Uhlen, M.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990  
A:Title: General colorimetric method for DNA diagnostics allowing direct solid-phase  
A:Reference number: A36043; MUID:90370827  
A:Accession: A36043  
A:Molecule type: DNA  
A:Residues: 294-402 <WAH>  
A:Experimental source: serotype L2  
R:Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.  
Gene 87, 105-112, 1990  
A:Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated tra  
A:Reference number: JQ0514; MUID:90236284  
A:Accession: A30472  
A:Molecule type: DNA  
A:Residues: 1-46;528-547 <LAM>  
A:Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:g144485  
R:Clarke, I.N.; Ward, M.E.; Lambden, P.R.  
Gene 71, 307-314, 1988  
A:Title: Molecular cloning and sequence analysis of a developmentally regulated cyste  
A:Reference number: JT0419; MUID:89138006  
A:Accession: JT0419  
A:Molecule type: DNA  
A:Residues: 30-547 <CLA>  
A:Cross-references: GB:M35148; NID:g144485; PIDN:AAA23119.1; PID:g144487  
A:Experimental source: serotype L1  
C:Genetics:  
A:Gene: omp2; omcB  
C:Function:  
A:Description: associated with differentiation of reticulate bodies into elementary b  
A:Note: essential for the structural integrity of the outer envelope of the elementar  
C:Superfamily: 60K cysteine-rich outer membrane protein  
C:Keywords: membrane protein; virulence  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-40/Domain: propeptide #status experimental <PRO>  
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MA

Query Match 72.2%; Score 2129.5; DB 1; Length 547;  
Best Local Similarity 70.7%; Pred. No. 3e-140;  
Matches 395; Conservative 77; Mismatches 72; Indels 15; Gaps 4;

QY 1 MSKLIRRVTVLALTSMAFCGAEAAVAESLITKIVASAEETKPAVPMTAKKVLVLR 60  
DB 1 MNKLIRRAVTITPAVTSVASLFSAGVLETSMAEFISTNVISLADTKAK--DNTSHKSKKAR 58  
QY 61 RN---KQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOQESCYGRLYSVKYNDDCNVEICQS 117  
DB 59 KHSKKEPVDK-----KKVAPVHESKA--TGPKQDCFCGRMYTVKVNDRNRVEITQA 108  
QY 118 VPEYATVGSYPPIELAILAIGCKDCVDVITQOLPCEAEFVSDDPETPTSDGKLWVKIDRL 177  
DB 109 VPKYATVGSYPPIELIITATGRKDCVDVITQOLPCEAEFVRSDEATPTADGKLWVKIDRL 168  
QY 178 GAGDKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVC 237  
DB 169 GQGEKSKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICVKOEGPENACLRCPVV 228  
QY 238 YKIEVNTVNGSAIARNVTVDNVPDGYSHASQORVLSFNGLDMRPGDKKVTVEFCPPORRG 297  
DB 229 YKINIVNQGTATARNVVENPVDGYAHSSGQVRVLTTLGDMQGEHRTITVEFCPLKRG 288  
QY 298 QITNVATVTCGGHKCSANVTTVVNEPCVOVNI SGADMSYVCKPVEYSISVSNPGDLVLH 357  
DB 289 RATNIAMVSVCGGHKNFASVTTVINEPCVQVSIAGADWSYVCKPVEYSISVSNPGDLVLR 348

Qy 358 DVVIQDTLPSPGVTVLEAPGGEICCNKVMIRKEMCPGETLQKLVKVAQVPGRETNQAV 417  
Db 349 DVVWDTLSPGVTVLEAAGAGACSCNKVWTVKELNPESLQKVLVRAQTGQCFNNVVV 408  
Qy 418 TSESCGCTCAETTTTHKGLAATHMCLVLTNDPICVGTNTYVYICVNTNRGSAEDTNS 477  
Db 409 KSCDCGCTCAEATTTWKGAATHMCLVLTNDPICVGTNTYVYICVNTNRGSAEDTNS 468  
Qy 478 LILFKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAI 537  
Db 469 LMLKFSKELQPVSGPTKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAI 528  
Qy 538 LSSDTLSPVSDTENTHY 556  
Db 529 LSSDTLTPVSDTENTHIY 547

RESULT 9  
F69009  
probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta H  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: F69009  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: F69009  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1474 <MTH>  
A:Cross-references: GB:AE000878; GB:AE000666; NID:g2622171; PIDN:AAB85563.1; PID:g262217  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1074  
C:Keywords: duplication

Query Match 5.1%; Score 151.5; DB 2; Length 1474;  
Best Local Similarity 24.3%; Pred. No. 0.019;  
Matches 83; Conservative 48; Mismatches 159; Indels 51; Gaps 10;  
Qy 236 VCYKIEVNTGSAIARNVTVDPDGYSHASQORVLSFNLGDMRGDKKVTVEFCPPOR 295  
Db 382 VRFTITVNYGPNATGVVYVTDLLPQLSFVSAS--ASRGITNTGTIWNLEYFETV 439  
Qy 296 RQGITNATVTCGGHKCSANTVTVVNEPCVQVNISGA-----DWS-----Y 337  
Db 440 TLNIT--ATVTATGAIVNANNVTGDFDPDMANNYASALNSPPASDLTIDKSVNPEPY 497  
Qy 338 VCKPVEYSISVSNPGLVLDHVDIOTLPSGVTVLEAPGGEICCNKVMIRKEMCPGETL 397  
Db 498 VGENIQYITVSNRCPDNaGVVDEVLPAGLIPISATPSKGSYNGTWNVGLNYLETA 557  
Qy 398 QFKLVKVAQVPGRETNQAVTS---ESNCGCTCAETTTTHKGL--AATHMCLVLTNDP- 452  
Db 558 TLTIIRYARNATGSLTNFANITSPNFDPNNDNDAEVV---GIPVADLLIVKQVSDPR 613  
Qy 453 ICVGTNYVYICVNTNRGSAEDTNSLILKFSKELQPIASSGPTKGTISGNTVVFDPALPKL 512  
Db 614 PDYGSVTVTVAVNLGPNATGVTVTDILSPGLVYL--SHVTVQCTYNATTGVN----- 666  
Qy 513 GSKESVEFSVTLKGIAPDARGEAILSDTLTSPVSDTENT 553  
Db 667 -----YIGALNYAASALMNLTVLVNTTGDNSNT 694

RESULT 10  
T28669  
surface protein 51c - Paramecium tetraurelia  
C:Species: Paramecium tetraurelia

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 07-Dec-1999  
C:Accession: T28669  
R:Nielsen, E.; You, Y.; Forney, J.  
J. Mol. Biol. 222, 835-841, 1991  
A:Title: Cysteine residue periodicity is a conserved structural feature of variable s  
A:Reference number: Z20504; MUID:92106337  
A:Accession: T28669  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-2233 <NIE>  
A:Cross-references: EMBL:M65164; NID:g159974; PID:g159975; PIDN:AAA61740.1  
C:Genetics:  
A:Genetic code: SOCS

Query Match 4.6%; Score 135; DB 2; Length 2233;  
Best Local Similarity 20.0%; Pred. No. 0.41; Mismatches 71; Indels 200; Gaps 34;  
Matches 125; Conservative 71; Mismatches 71; Indels 200; Gaps 34;  
Qy 5 IRRVVTVLALTSMSASFASGGIEAAVAESLI-----TKIVASAEFTKPAPV 49  
Db 1 MKRLLLIAMISIAATCOVWSKSEACTCAQLTSGDCARNSNCNNTTKLACEVPGSTGPV 60  
Qy 50 PMTAKKVLVRRNKQPVQKSRGAFCDKEFYPCBEGRCOPVEAOQESCYG----- 99  
Db 61 TVT-----KNYKSLYC-----EGLAQTDCLKLNCAWIDNKCCTFTS 98  
Qy 100 RLYSVKVNDDC-----NVEIQSVPEVATVGSVPYIE-----LAIGKK 138  
Db 99 CTPYEKTIKDCOAISKRCITDGTICVEIDLCT---YLTSTCYONKAGNYCVWDETAK 155  
Qy 139 DCVDDVVTQOLP-----CEA--EFVSDPETTPSDGKLVNWKIDRLGAGDKCKITVW 188  
Db 156 KCSDDTECAQLPALTKDSECRAYLKF-----ECTAKPAGGCV-----DSGTNCADQVS 204  
Qy 189 VKPLKEGC-----CFTAATVCAPELRSVTKCOGPAICIKOEGPDCACLRCPVCYK 239  
Db 205 V-----EGCVTNKTRSVNCFWDTTNKCFD---KKENASTNKTHT--VDCQAF--LP7C-- 252  
Qy 240 IEVNTGSAIARNVTVDPDGYSHASQORVLSFNLGDMRGDK-----KVFTVEFC 292  
Db 253 -TAKDGGCVDIKTCADGKIKECKIDSAKKECYWSDKDLCKDKIKASAPNTLTNNSDC 311  
Qy 293 POR--RQGITNATV---TYCGGHKCSANTVTVVNEPCV--OVNIGADWSYCKPVEYS 345  
Db 312 QKQFLASCITNGAGCVDDTSCG-----SSVQEQCAVNRNNRECTWNGSCK-----D 359  
Qy 346 ISVSNPG--DLVLHD-----VVI-----QDTLPSPGV 370  
Db 360 KTCENAGTDIVGHDOCSYKACTGKANNAGGCKRSCDAPTTIISNAGCEDYLPNGKC 419  
Qy 371 VLEAPGGEI---CCNKVWVR---IKEM---CPGETLQKLVK--AQVPGRETNQAVT 418  
Db 420 IAKDGGCISNTTCSAILLKDACVQDNKKQCYWDTVGNCLDKTCATLPTRLNLSHLCN 479  
Qy 419 SESNCGTCT-----SCAETTTTHKGLAATHMCLVLTNDPICVGTNYVYICVNTNRGSAED 473  
Db 480 GEIN--TCTVSSSGTCVDLLC---ENVVDKNCVKDKSGADCV---YYGSCYQKQCSAAS 531  
Qy 474 TNVSLILKFSKELQPIASSGPTKGTI 499  
Db 532 QDNTTHAQOEYLPACTLSNTKKGCI 557

RESULT 11  
D75622  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: D75622  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J  
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: D75622  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-756 <WHI>  
A:Cross-references: GB:AE001826; PID:96460827; PIDN:AAF12630.1; PID:96460926; TIGR:DRB00  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRB0039  
A:Map position: megaplasmid  
A:Genome: plasmid  
A:Note: plasmid MPI

Query Match 4.4% Score 130.5; DB 2; Length 756;  
Best Local Similarity 22.6% Pred. No. 0.25;  
Matches 121; Conservative 68; Mismatches 217; Indels 129; Gaps 31;  
QY 108 DDCNV-EICQSVPEYATGSPYPIEILAIKGGKDCVGVVIT-----QOL 149  
DB 23 DDNNVAQLVATPPHLELRKAFGAERLRPGDGTGVTATNTAGAAAPDVVLEDPGLGOOL 82  
QY 150 PCEAEFVSSDPTTSDGKLWKID-----RLGAGDKCKIT--V 187  
DB 83 DAGLAV---PGSARTDRGLEYSADGASQAAEPAAVRGVRVQAGTLEPGEQTTLTFRM 139  
QY 188 WKPLKEGCGF-TAATVCACP-----ELR-----SYTKGCPAICIKQEGPDC-- 229  
DB 140 EARPSAENLLNNVATSVSVTGEQAQSDTLQVFLPGVAGLPQGP---LAPECTAAGD 196  
QY 230 -----ACLRCPVCYKTEVNTGSA-----IARNVTVDNPPDGYSHASQORVLSNLGDM 279  
DB 197 QTLVAVTGPACFDHTVQNTQDMERLSVSTQGOATPOLLGAAGPLPQPETLA-- 254  
QY 280 RPDGKKVTFVEFCQRRGQITNVATVYTCGGHKCSANVT-----IVNNEPCVQVNIS-- 331  
DB 255 -PGEOR--QVRVCYDLRSQAQPLAQV--IAGERGTSNATDLRRVETQRPGLRKTVSKV 310  
QY 332 GA-DW---SYVCK--PVEYSISVSNPGDLVLHDVVIQDTLPSPGVTVLEAPG-----GEIC 380  
DB 311 GAPDAPGSAVTSDELEYLTSLVNPVQAQPLAGVQVLDPLPAGTFEVSASDGGALLGAAS 370  
QY 381 CNKYVWRIKEMCPGETIQFLVKAQVPGR-----FTNOVAVTSNCGTCTSCAETT 433  
DB 371 AAQVANTLGLDLPAGATRTLTLRVRV---GRDVRDQELRNWFELTSSLEPAPLHNSAASA 427  
QY 434 THWGLAATHMCVLDTNDPICVGVNTVYRICVTNRGSAEDTNVSLI-----LKF--- 482  
DB 428 VVW-NTAPLLSKTLDRD--AAPGDLLTYTLTKN-PSASTALVDLVTITDTPAAALKYVTG 484  
QY 483 SKELOPIASSGPTKGTISGNTVVDALPKLSKESVEFVTLKGIAPDARGEAI 537  
DB 485 TSRLAGVPTADPAE---NGGELW-RVPRLAGESLTLSYGLR-VLPQ-AGELL 533

RESULT 12  
S16148  
gene serrate protein precursor - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 31-Dec-1991 #sequence\_revision 02-Aug-1994 #text\_change 17-Nov-2000  
C:Accession: S16148; S16878; A36666  
R:Thomas, U.; Speicher, S. A.; Knust, E.  
A:Title: The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a co  
Development 111, 749-761, 1991  
A:Reference number: S16148; MUID:91347903  
A:Accession: S16148  
A:Molecule type: mRNA  
A:Residues: 1-1408 <THOI>  
A:Cross-references: EMBL:X56811  
R:Thomas, U.  
submitted to the EMBL Data Library, November 1990

A:Reference number: S16878  
A:Accession: S16878  
A:Molecule type: mRNA  
A:Residues: 1-1351, 'T', 1353-1408 <THO2>  
A:Cross-references: EMBL:X56811; NID:98563; PID:98564  
R:Fleming, R. J.; Scottgale, T. N.; Diederich, R. J.; Artavanis-Tsakonas, S.  
Genes Dev. 4, 2188-2201, 1990  
A:Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential  
A:Reference number: A36666; MUID:91099666  
A:Accession: A36666  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-15, 20-26, 'A', 28-1408 <FLE>  
A:Cross-references: GB:M35759; NID:9158605; PID:9158606  
C:Genetics:  
A:Gene: FlyBase:Ser  
A:Cross-references: FlyBase:FBgn0004197  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: glycoprotein; transmembrane protein  
F:1-84/Domain: signal sequence #status predicted <SIG>  
F:85-1408/product: gene serrate protein #status predicted <MAT>  
F:85-1221/Domain: extracellular #status predicted <EXT>  
F:283-316/Domain: EGF homology <EG01>  
F:319-348/Domain: EGF homology <EG02>  
F:355-388/Domain: EGF homology <EG03>  
F:395-488/Domain: EGF homology #status atypical <EG04>  
F:495-526/Domain: EGF homology #status atypical <EG05>  
F:533-608/Domain: EGF homology #status atypical <EG06>  
F:615-645/Domain: EGF homology <EG07>  
F:652-683/Domain: EGF homology <EG08>  
F:690-720/Domain: EGF homology #status atypical <EG09>  
F:727-796/Domain: EGF homology <EG11>  
F:803-834/Domain: EGF homology <EG12>  
F:841-876/Domain: EGF homology <EG13>  
F:883-914/Domain: EGF homology <EG14>  
F:921-1060/Region: cysteine-rich  
F:997-1246/Domain: transmembrane #status predicted <TML>  
F:1247-1408/Domain: intracellular #status predicted <INT>  
F:152,196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (

Query Match 4.4% Score 130.5; DB 2; Length 1408;  
Best Local Similarity 20.5% Pred. No. 0.51;  
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;  
QY 14 LTSMSCFPASGGIEAAVAESLTIKIVASAEITKPAVPMTAKVRLVRRNKQVEKSRGA 73  
DB 547 LTTTATAITGNSJSSVALLAALTSVASTSLAIGPC-INAKECR-----NQP-----GS 594  
QY 74 FCDKEFVPCBEGRCQPVEAQOQSCYGRLYSVKVNDCNVEICQSVPEYATGSPYPIEIL 133  
DB 595 FA---CICKEG-----WGVTCENLDDC-VGQCRN----- 621  
QY 134 AIGKKDCVDTTQQLPCEAFVSSDPE-----TTPTSDG-----KLVMK-----IDRLG- 178  
DB 622 ---GATCIDLVNDYRCACASGFTGRDCEIDIDECATSPCRNGGECVDMVGKFCICPLGY 678  
QY 179 AGDKC---KITVWVKPLKEGCGCTA---ATVCAPELRSYTKCGQ-PAIC----- 221  
DB 679 SGSLCEAEKENCPTSPCLGHECLNTPEGYCHCPCPPDRAGKHQELRPLCSOPPCNEGCFA 738  
QY 222 -----IKQEGPDC-----AKLRCPV---CYKIEVVTGSAIA 250  
DB 739 NVSLATSATTTTTTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGTGTFC 797  
QY 251 RNVTVDNVDP-----DG-----YSHASQORVLS-----FNLGDMRPG-- 282  
DB 798 HNLNECSPNCRNGGICLDGDFTCCECMGWTGKRCSEATGCGYAGOCQNGGTGMPGAP 857  
QY 283 DKKV-----FTVEFCQRRGQITNVATVYTCGGHKCSANVTTVVNEPCVQVNISGA 333  
DB 858 DKALOPHCRCAPGWTGLFCAE-----AIDQCRGQFCHNGGT-----CE-----SGA 898

RESULT 14  
C75489  
conserved hypothetical protein - Deinococcus radiodurans (strain RI)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: C75489  
R:White, O.; Eisen, J. A.; Heidelberg, J. F.; Hickey, E. K.; Peterson, J. D.; De  
S.: Shen, M.; Vamathevan, J. J.; Lam, P.; McDonald, L.; Utterback, T.; Zall  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiod  
A:Reference number: AY5250; MUID:20036896

A:Cross-references: CB:AE001925; CB:AE005513; NID:G6458383; PIDN:AAF10262.1;  
A:Experimental source: strain R1  
C:Genetics;  
A:Gene: DR0686  
A:Map position: 1

[illegible]

RESULT 15  
T25878 hypothetical protein T10E9.4 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T25878  
R:Greco, T.; Hawkins, M.  
submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of *C. elegans* cosmid T10E9.



Tue May 28 08:55:19 2002

A Reference number: Z20104  
A Accession: T25878  
A Status: preliminary; translated from GB/EMBL/DBJ  
A Molecule type: DNA  
A Residues: 1-693 <GRE>  
A Cross-references: EMBL:U97403; PIDN:AAB52472.1; GSPDB:GN00019; CESP:T10E9.4  
A Experimental source: strain Bristol N2; clone T10E9  
C Genetics:  
A Gene: CESP:T10E9.4  
A Map position: 1  
A Introns: 16/1; 40/1; 87/1; 136/1; 185/1; 264/1; 322/1; 488/1; 565/2

Query Match 4.3%; Score 127; DB 2; Length 693;  
Best Local Similarity 19.8%; Pred. No. 0.4; Indels 222; Gaps 26;  
Matches 103; Conservative 41; Mismatches 41; Gaps 26;

QY 46 PAPVMTAKKVLV-----RNKQPVQKS-----RGAFCDKRYFPCEEG 85  
DB 185 PIPFSFTLKNNSNVPLTWAEEGLEKSKANNQSYKVKTCTTNSGCGDNFCDSVYNPTFN 244  
QY 86 -----RCQVEAQEQSCYGR-----LYSVKV-----NDDC--NVEIC--QSVPEY 121  
DB 245 HRFDRSGPKSSLOFCYKQPDLPDLKHVKVDLGLKLCHYNEDCGNETEICHIDNIIVY 304  
QY 122 ATVGSYPPIEILAIKKDCVDVITQO-----LPCAEFVSDDPETTPTSDGKLVM 172  
DB 305 ANVTOPEKIAVPGI-----CVHVDVCSAADDEKTHAVLPVNSQFC'KEDMHCQ----- 351  
QY 173 KIDRLGAGDKCKIT-----VWVKPLKEGCECTAATVCAPBELRSYTKCGQPAICIKOE 225  
DB 352 -----NAGVSANOTEXSHHCRSYSDPNKVCCEKPKCKHGAETQSPVANLMQCKYE 406  
QY 226 GPDCAICLRCPVCYKIEVVNTGSAIARNVTDNPVDGYSHASQGVLSFNLGDMRPGDKK 285  
DB 407 --DCA-----GDLK--SDKK 417  
QY 286 VFTVEFCPQRRGQITNV-----ATVTCGGHKCSANVTIVNEP-CVQVNI SGADMSYV 338  
DB 418 RLEL-WCEK-----VNCCKDIGSTVTKDKHGTCLDYATPLYNEPKCDV--VEGKSSGV 471  
QY 339 CKPVEYSISVSNPGDLVLHDVVQDTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETL- 397  
DB 472 CK-----TKGG-----VCREGHCCP'SLTLT 491  
QY 398 -----QFKLVVKAQVQGRF-----TNOVAVTSESNCGTC 426  
DB 492 IAPSGNGTESATPTLGPYPYLTNYPDCANKPIPSQFSTYAFCDPDTRVGI LGRHL--- 548  
QY 427 TSCAETTHWKGLAATHM--CVLDTNDPTCYGENTVYRIC 464  
DB 549 --TGBERTEVKGSGACSSNNKCKSGT---VCVYVNI NKHVC 583

Search completed: May 25, 2002, 22:16:36  
Job time: 5270 sec



P29579 methanobact  
Q28295 canis famil  
P98092 bombyx mori  
O14246 homo sapien  
Q9tv36 sus scrofa  
P34576 caenorhabdi  
Q07008 rattus norv  
Q05793 mus musculu  
P98160 homo sapien  
Q02817 homo sapien  
O62918 rattus norv  
P39712 saccharomyc

34 104.5 3.5 499 1 YP25\_METTE  
35 104.5 3.5 2813 1 VWF\_CANFA  
36 104.5 3.5 3133 1 HMCT\_BOMMO  
37 104 3.5 886 1 EMRI\_HUMAN  
38 104 3.5 2871 1 FBNI\_PIG  
39 104 3.5 3051 1 YNX3\_CAEEL  
40 103 3.5 2531 1 NTC1\_RAT  
41 102.5 3.5 3707 1 PGBM\_MOUSE  
42 102.5 3.5 4393 1 PGBM\_HUMAN  
43 102.5 3.5 5179 1 MUC2\_HUMAN  
44 102 3.5 816 1 NEL2\_RAT  
45 102 3.5 1322 1 YAG3\_YEAST

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

DM protein - protein search, using sw model

Run on: May 25, 2002, 22:16:42 ; Search time 39.38 Seconds  
(without alignments)  
546.675 Million cell updates/sec

Title: US-09-523-647-2

Perfect score: 2951

Sequence: 1 MSKLIRRVTVLALTSASC.....ILSDTLTSPVSDTENTHYV 556

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID        | Description        |
|------------|--------|-------------|--------|--------------|--------------------|
| 1          | 2951   | 100.0       | 556    | 1 OM6_CHLPN  | P23700 chlamydia p |
| 2          | 2544.5 | 86.2        | 557    | 1 OM6_CHLPS  | P23701 chlamydia p |
| 3          | 2162.5 | 73.3        | 547    | 1 OM6C_CHLTR | P23603 chlamydia t |
| 4          | 2161.5 | 73.2        | 547    | 1 OM6C_CHLTR | P26758 chlamydia t |
| 5          | 2160.5 | 73.2        | 547    | 1 OM6D_CHLTR | P18151 chlamydia t |
| 6          | 2129.5 | 72.2        | 547    | 1 OM6L_CHLTR | P21354 chlamydia t |
| 7          | 130.5  | 4.4         | 1408   | 1 SERR_DROME | P18168 drosophila  |
| 8          | 129    | 4.4         | 772    | 1 ANNU_SCHAM | P52183 schistocerc |
| 9          | 123.5  | 4.2         | 2139   | 1 CRB_DROME  | P10040 drosophila  |
| 10         | 123    | 4.2         | 1295   | 1 GLP1_CAEEL | P13508 caenorhabdi |
| 11         | 115.5  | 3.9         | 577    | 1 TRBM_MOUSE | P15306 mus musculu |
| 12         | 115.5  | 3.9         | 2911   | 1 FBNI_HUMAN | P35556 homo sapien |
| 13         | 113.5  | 3.8         | 555    | 1 DP87_DICDI | Q04503 dictyosteli |
| 14         | 113.5  | 3.8         | 1808   | 1 TENA_CHICK | P10039 gallus gall |
| 15         | 113    | 3.8         | 2807   | 1 FBNI_MOUSE | Q61555 mus musculu |
| 16         | 112    | 3.8         | 2871   | 1 FBNI_MOUSE | Q29116 sus scrofa  |
| 17         | 111.5  | 3.8         | 1746   | 1 TENA_PIG   | P24821 homo sapien |
| 18         | 111    | 3.8         | 2201   | 1 FBNI_HUMAN | P35555 homo sapien |
| 19         | 110.5  | 3.7         | 2871   | 1 FBNI_HUMAN | Q03610 caenorhabdi |
| 20         | 110    | 3.7         | 1416   | 1 YNBI_CAEEL | P07207 drosophila  |
| 21         | 110    | 3.7         | 2703   | 1 NOTC_DROME | P98133 bos taurus  |
| 22         | 109.5  | 3.7         | 2871   | 1 FBNI_BOVIN | Q99088 xenopus lae |
| 23         | 109    | 3.7         | 892    | 1 LD12_XENLA | P17053 paramecium  |
| 24         | 108    | 3.7         | 2704   | 1 G168_PARPR | Q9uh18 homo sapien |
| 25         | 107    | 3.6         | 967    | 1 ATSI_HUMAN | Q03376 chironomus  |
| 26         | 107    | 3.6         | 1700   | 1 BAR3_CHITE | P98165 gallus gall |
| 27         | 106.5  | 3.6         | 863    | 1 LDVR_CHICK | P03517 punta toro  |
| 28         | 106.5  | 3.6         | 1313   | 1 VGLM_PTPV  | P98164 homo sapien |
| 29         | 106.5  | 3.6         | 4655   | 1 LRP2_HUMAN | P34504 caenorhabdi |
| 30         | 106    | 3.6         | 1246   | 1 YMV2_CAEEL | P30432 drosophila  |
| 31         | 105.5  | 3.6         | 1680   | 1 FUR2_DROME | Q9wuc1 rattus norv |
| 32         | 105    | 3.6         | 967    | 1 ATSI_RAT   | P46531 homo sapien |
| 33         | 105    | 3.6         | 2444   | 1 NTC1_HUMAN |                    |

ALIGNMENTS

RESULT 1

| ID | OM6_CHLPN   | STANDARD; | PRT; | 556 AA. |
|----|---|-----------|------|---------|
| AC | P23700; Q9JOI6;   |           |      |         |
| DT | 01-NOV-1991 (Rel. 20, Created)  |           |      |         |
| DT | 01-NOV-1991 (Rel. 20, Last sequence update)   |           |      |         |
| DE | 16-OCT-2001 (Rel. 40, Last annotation update)   |           |      |         |
| DE | 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).      |           |      |         |
| DE | OMCB OR OMP2 OR CPN0557 OR CP0195.  |           |      |         |
| GN | Chlamydia pneumoniae (Chlamydia pneumoniae).  |           |      |         |
| OS | Chlamydia pneumoniae (Chlamydia pneumoniae).  |           |      |         |
| OC | Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.   |           |      |         |
| OX | NCBI_TaxID=83558;   |           |      |         |
| RN | [1]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN=IOL-207;   |           |      |         |
| RC | MEDLINE=90384850; PubMed=2402463;   |           |      |         |
| RA | Watson M.W., Al-Nahdawi S., Lamden P.R., Clarke I.N.;   |           |      |         |
| RT | "The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207."; |           |      |         |
| RL | Nucleic Acids Res. 18:5299-5299(1990).  |           |      |         |
| RN | [2]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN=CWL029;  |           |      |         |
| RC | MEDLINE=99206606; PubMed=10192388;  |           |      |         |
| RA | Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  |           |      |         |
| RA | Olinger L., Grimwood J., Davis R.W., Stephens R.S.;   |           |      |         |
| RT | "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";  |           |      |         |
| RL | Nat. Genet. 21:385-389(1999).   |           |      |         |
| RN | [3]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN=AR39;  |           |      |         |
| RC | MEDLINE=20350255; PubMed=10684935;  |           |      |         |
| RA | Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,  |           |      |         |
| RA | White O., Hickey E.K., Peterson J., Utterback T., Berry K., Dodson R.,  |           |      |         |
| RA | Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  |           |      |         |
| RA | Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  |           |      |         |
| RA | Risen J., Fraser C.M.;  |           |      |         |
| RT | "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AK39.";                                      |           |      |         |
| RL | Nucleic Acids Res. 28:1397-1406(2000).  |           |      |         |
| RN | [4]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN=J138;  |           |      |         |
| RC | MEDLINE=20330349; PubMed=10871362;  |           |      |         |
| RA | Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  |           |      |         |
| RA | Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;   |           |      |         |
| RT | "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";                  |           |      |         |
| RL | Nucleic Acids Res. 28:2311-2314(2000).  |           |      |         |
| RN | [5]   |           |      |         |
| RP | SEQUENCE FROM N.A.  |           |      |         |
| RC | STRAIN=J138;  |           |      |         |
| RC | Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,  |           |      |         |
| RA | Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,  |           |      |         |
| RA | Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,   |           |      |         |

RA Ishii K., Shiba T., Hattori M., Kuhara S.;  
RT "Comparison of outer membrane protein genes omp and pmp in the whole  
RT US";  
RL genome sequences of Chlamydia pneumoniae isolates from Japan and  
CC Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES  
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL  
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT  
CC VIRULENCE FACTOR.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53511; CAA37590.1; -;  
DR EMBL; AE001640; AAD18697.1; -;  
DR EMBL; AE002180; AAF38068.1; -;  
DR EMBL; AP002547; BAA98763.1; -;  
DR EMBL; AB033786; BAA85939.1; -;  
DR PIR; S12602; S12602.  
DR PHCI-2DPAGE; P23700; -;  
DR TIGR; CP0195; -;  
DR InterPro; IPR003506; Chlam\_OMP6.  
DR PRINTS; PR01336; CHLAMIDIAOM6.  
DR Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.  
KW SIGNAL 1 22  
FT PROPEP 23 40 POTENTIAL.  
FT CHAIN 41 556 60 KDA OUTER MEMBRANE PROTEIN.  
SQ SEQUENCE 556 AA; 59719 MW; 8D7ED9234CC99458 CRC64;

Query Match 100.0%; Score 2951; DB 1; Length 556;  
Best Local Similarity 100.0%; Pred. No. 1.5e-213;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKLIIRVTVTLATSWASCFASGIEAAVAESLITKIVASAEKTPAPVPMATAKRVLRV 60  
DB 1 MSKLIIRVTVTLATSWASCFASGIEAAVAESLITKIVASAEKTPAPVPMATAKRVLRV 60  
QY 61 RNKQPVQKSRGAFCDKEFFVCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120  
DB 61 RNKQPVQKSRGAFCDKEFFVCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120  
QY 121 YATVGSYPYIEILAIKDKCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180  
DB 121 YATVGSYPYIEILAIKDKCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180  
QY 181 DKCKITVWKPLKGGCCFTATVACAPCLRSYTKGQPAICIKOEGDPCACLRCPVCYKI 240  
DB 181 DKCKITVWKPLKGGCCFTATVACAPCLRSYTKGQPAICIKOEGDPCACLRCPVCYKI 240  
QY 241 EYVNTGSATARNVTVDNVPDGYSHASQGVLSFNLGDMRPGDKKFTVEFCPPORRGQIT 300  
DB 241 EYVNTGSATARNVTVDNVPDGYSHASQGVLSFNLGDMRPGDKKFTVEFCPPORRGQIT 300  
QY 301 NVATVTCGCGHKCSANVTTVNPEPCVQVNTSGADWSVYCKPVEYSISVSNPGLVLHDV 360  
DB 301 NVATVTCGCGHKCSANVTTVNPEPCVQVNTSGADWSVYCKPVEYSISVSNPGLVLHDV 360  
QY 361 IQDTLPVSGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVVKAQVPGFTNOVATSE 420  
DB 361 IQDTLPVSGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVVKAQVPGFTNOVATSE 420  
QY 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPFCVGTNTYRICVTVNRGSAEDTNVSLIL 480  
DB 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPFCVGTNTYRICVTVNRGSAEDTNVSLIL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540

Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHVV 556  
DB 541 DTLTSPVSDTENTHVV 556  
RESULT 2  
OM6\_CHLPS  
ID OM6\_CHLPS STANDARD; PRT; 557 AA.  
AC P23701;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane  
DE protein) (CRP) (60 kDa cysteine-rich OMP).  
DE OMCB OR OMP2 OR ENVB.  
OS Chlamydia psittaci (Chlamydia phila psittaci).  
OX Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
OX NCBI\_TaxID=83554;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EAE/A22/M;  
RX MEDLINE=90384851; PubMed=2402464;  
RT Watson M.W.; Lambden P.R.; Clarke I.N.;  
RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane  
RT protein of Chlamydia psittaci strain EAE/A22/M.";  
RL Nucleic Acids Res. 18:5300-5300(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=6BC;  
RX MEDLINE=91267949; PubMed=2050637;  
RA Everett K.D.E.; Hatch T.P.;  
RT "Sequence analysis and lipid modification of the cysteine-rich  
RT envelope proteins of Chlamydia psittaci 6BC.";  
RL J. Bacteriol. 173:3821-3830(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Watson M.W.;  
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES  
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL  
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT  
CC VIRULENCE FACTOR.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53512; CAA37592.1; -;  
DR EMBL; M61116; AAB61619.1; -;  
DR PIR; S12603; S12603.  
DR InterPro; IPR003506; Chlam\_OMP6.  
DR PRINTS; PR01336; CHLAMIDIAOM6.  
DR Outer membrane; Transmembrane; Signal; Virulence.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 40 POTENTIAL.  
FT CHAIN 41 557 60 KDA OUTER MEMBRANE PROTEIN.  
FT VARIANT 45 45 A -> S (IN STRAIN 6BC).  
FT VARIANT 73 73 E -> G (IN STRAIN 6BC).  
SQ SEQUENCE 557 AA; 59843 MW; 0D444F09EAA073C6 CRC64;

Query Match 86.2%; Score 2544.5; DB 1; Length 557;  
Best Local Similarity 85.1%; Pred. No. 4e-183;  
Matches 474; Conservative 30; Mismatches 52; Indels 1; Gaps 1;

Tue May 28 08:55:20 2002

QY 1 MSLKLRVVTVLALTSMAFCASGIEAARVAESLITKIVASAEKTPAPV-PMTAKKRVLV 59  
DB 1 MSLKLRVVTVLALTSMAFCASGIEAARVAESLITKIVASAEKTPAPV-PMTAKKRVLV 60  
QY 60 RNNKQPVQKSGAFCDKEFYCEBGRCPVPAQAQESCYGRILYSVKVNDNCNVEICQSV 119  
DB 61 RNNKQPVQKSGAFCDKEFYCEBGRCPVPAQAQESCYGRILYSVKVNDNCNVEICQSV 120  
QY 120 EYATVGSPIELAIAGKDCVDVITQQLPCEAEFVSDPPTPSDGLVWKIDRLGA 179  
DB 121 EYATVGSPIELAIAGKDCVDVITQQLPCEAEFVSDPPTPSDGLVWKIDRLGA 180  
QY 180 GOKCKITVWVKPLKEGCEFTAAATVACAPELRSYTKCGQPAICIKQEGPCACLRCPVCYK 239  
DB 181 GOKCKITVWVKPLKEGCEFTAAATVACAPELRSYTKCGQPAICIKQEGPCACLRCPVCYK 240  
QY 240 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 299  
DB 241 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 300  
QY 300 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 359  
DB 301 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 360  
QY 360 VIQDTLPSGVTVLEAPGCEICCNKVRIRKEMCPGETLQFLKLVKVAQVGRFTNOVAVTS 419  
DB 361 VIQDTLPSGVTVLEAPGCEICCNKVRIRKEMCPGETLQFLKLVKVAQVGRFTNOVAVTS 420  
QY 420 ESNCGTCSAEFTTHWKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 479  
DB 421 NSDCGTCSAEFTTHWKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 480  
QY 480 LKFSKELQPIASSGPTKTGTSIGNTVVFDALPKLGSKEVSFVTLKGIAPDARGAELIS 539  
DB 481 LKFSKELQPIASSGPTKTGTSIGNTVVFDALPKLGSKEVSFVTLKGIAPDARGAELIS 540  
QY 540 SDTLTSPVSDTENTHVV 556  
DB 541 SDTLTSPVSDTENTHVV 557

## RESULT 3

ID OM6E\_CHLTR STANDARD; PRT; 547 AA.  
AC P23603;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 60 kDa outer membrane protein, serovar E precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).  
GN Omp2 OR Omp2B.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DK 20 / SEROVAR E;  
RX MEDLINE=91067486; PubMed=2251143;  
RA Coles A.M., Allan I., Pearce J.H.;  
RT "The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis serovar E.";  
RL Nucleic Acids Res. 18:6713-6713(1990).  
RP SEQUENCE FROM N.A.  
RC STRAIN=BOUR / SEROVAR E;  
RX MEDLINE=91147205; PubMed=1997423;  
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;  
RT "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";  
RL Infect. Immun. 59:1196-1201(1991).  
RN [3]

RP SEQUENCE FROM N.A.  
RA Zhang Y.X., Caldwell H.D.;  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
CC FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.  
CC SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.  
CC MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAIN DK 20.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: X55903; CAA39396.1; -;  
CC EMBL: X54389; CAA38259.1; -;  
CC EMBL: M85196; AAA23154.1; -;  
CC PIR: S13120; S13120. Chlam\_OMP6..  
CC InterPro: IPR003506; Chlam\_OMP6..  
CC PRINTS: PR01336; CHLAMIDIAOM6. Signal; Virulence.  
CC Outer membrane; Transmembrane; Signal; Virulence.  
CC SIGNAL 1 22 POTENTIAL.  
CC PROPEP 23 40  
CC CHAIN 41 547  
CC VARIANT 33 34  
CC VARIANT 121 121  
CC VARIANT 132 132  
CC VARIANT 458 458  
CC SEQUENCE 547 AA; 58708 MW; 052066084F4E20AB CRC64;

Query Match 73.3%; Score 2162.5; DB 1; Length 547;  
Best Local Similarity 72.4%; Pred. No. 1.5e-154;  
Matches 403; Conservative 72; Mismatches 71; Indels 11; Gaps 4;

QY 1 MSLKLRVVTVLALTSMAFCASGIEAARVAESLITKIVASAEKTPAPV-PMTAKKRVLV 60  
DB 1 MSLKLRVVTVLALTSMAFCASGIEAARVAESLITKIVASAEKTPAPV-PMTAKKRVLV 58  
QY 61 RNNKQPVQKSGAFCDKEFYCEBGRCPVPAQAQESCYGRILYSVKVNDNCNVEICQSV 119  
DB 59 K-----HSKETLVDKKEVAPVHESKA--TGPQKQSCFGRMYTKVNDNRNVEITQAVP 110  
QY 120 EYATVGSPIELAIAGKDCVDVITQQLPCEAEFVSDPPTPSDGLVWKIDRLGA 179  
DB 111 EYATVGSPIELAIAGKDCVDVITQQLPCEAEFVSDPPTPSDGLVWKIDRLGA 170  
QY 180 GOKCKITVWVKPLKEGCEFTAAATVACAPELRSYTKCGQPAICIKQEGPCACLRCPVCYK 239  
DB 171 GEKSKITVWVKPLKEGCEFTAAATVACAPELRSYTKCGQPAICIKQEGPCACLRCPVCYK 230  
QY 240 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 299  
DB 231 INVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 290  
QY 300 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 359  
DB 291 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 350  
QY 360 VIQDTLPSGVTVLEAPGCEICCNKVRIRKEMCPGETLQFLKLVKVAQVGRFTNOVAVTS 419  
DB 351 VIEDTLPSPGVTVLEAPGCEICCNKVRIRKEMCPGETLQFLKLVKVAQVGRFTNOVAVTS 410  
QY 420 ESNCGTCSAEFTTHWKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 479  
DB 411 NSDCGTCSAEFTTHWKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 470  
QY 480 LKFSKELQPIASSGPTKTGTSIGNTVVFDALPKLGSKEVSFVTLKGIAPDARGAELIS 539  
DB 481 LKFSKELQPIASSGPTKTGTSIGNTVVFDALPKLGSKEVSFVTLKGIAPDARGAELIS 539

```
Db 471 LKFSKELQPVFSFGPTKGTITGNTVTVFDSLPLRGLSKETVEFSVTLKAVSAGDARGEAILS 530
QY 540 SDTLTPSPVSDTENTHIVY 556
Db 531 SDTLTPVPSDENTHIIY 547

RESULT 4
OM6D_CHLTR STANDARD; PRT; 547 AA.
AC P26758;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovar C precursor (Cysteine-rich
DE outer membrane protein) (60-kDa Crp).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-3 / SEROVAR C;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
RA Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative
RT 15-kilodalton protein between the trachoma and lymphogranuloma
RT venereum biovars of Chlamydia trachomatis.";
RL Infect. Immun. 59:1196-1201(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Zhang Y.X., Caldwell H.D.;
RA Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54388; CAA38257.1; -
DR EMBL; M85197; AAB23159.1; -
DR PIR; C43584; C43584.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22
FT PROPEP 23 40 POTENTIAL.
FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR C.
SQ SEQUENCE 547 AA; 58680 MW; 817BA5DC7FEA65D4 CRC64;

Query Match
Best Local Similarity 73.2%; Score 2161.5; DB 1; Length 547;
Matches 402; Conservative 74; Mismatches 68; Indels 15; Gaps 4;

QY 1 MSLKLRVVTVLALTSMAISCFASGIEAAVSLTKIVASAEKTPAPVPMATKKVRLVR 60
Db 1 MNKLIRRAVTITFATISVASLFSAGVLETSMAESLSTNVISLADTKAK--DNTSHKSKAR 58
QY 61 RN---KQPEVKSRGAFCDKFEYCEECRCQCPVQAQSCYGRLYSVKVNDCNVVEIQCS 117
Db 59 KNHSEKTPVDR-----KEVAPVHESKA--TGPKQDSCEGRMYTVKVNDRNVEITQA 108
QY 118 VPEYATVGSPIPIELAILGKKKDCVDVITQQLPCEAEFVSRDPTTPTSDGKLWVKIDRL 177
```

```
Db 109 VPEYATVGSPIPIELAITATKRCDCVDVITQQLPCEAEFVSRDPTTPTADGKLWVKIDRL 168
QY 178 GAGDKCKITVWKPLKEGCCFTAAATVCACPELRSYTKCGOPATICKQEGPDCACLRCPVC 237
Db 169 GQGEKSKITVWKPLKEGCCFTAAATVCACPEIRSVTKCGOPALCVKQEGPENACLRCPVV 228
QY 238 YKIEVYNTGSATARNVTDNPVDPGYSHASGORVLSFNLDGMRPDKKVFTVEFCPORRG 297
Db 229 YKINVVNOGTATARNVVENPVPDGYAHSSGORVLTFTGLDQPGQGEHRITITVEFCPLKRG 288
QY 298 QITNATVYTCGGHKCSANVTTVNPECVQVNIISGADWSYCKPVPYSISVSNPGLDLVLR 357
Db 289 RATNIATVSYCGGHKNTASVTTVINEPCVOVSIAGADWSYCKPVEYVLSVSNPGLDLVLR 348
QY 358 DVVIQDTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQPKLVVKAQVPGRTNOVAV 417
Db 349 DVVVEDTLSPGVTVLEAAGAQISCNKVVWTVKELNPGESLOYKVLVRAOTPGQFTNNVVV 408
QY 418 TSESNCGTCTSCAETTTTHWKGIAATHMVCVLDNDPFCVGENTVYRICVTNRGSAEDTNVS 477
Db 409 KSCSDCGTCTSCAETTVWKGVAATHMVCVVDTCDPVCVGVNTVYRICVTNRGSAEDTNVS 468
QY 478 LILKFSKELQPIASSGPTKGTISGNTVTVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAI 537
Db 469 LMLKFSKELQPVFSFGPTKGTITGNTVTVFDSLPLRGLSKETVEFSVTLKAVSAGDARGEAI 528
QY 538 LSSDTLTPSPVSDTENTHIVY 556
Db 529 LSSDTLTPVPSDENTHIIY 547

RESULT 5
OM6D_CHLTR STANDARD; PRT; 547 AA.
AC P18151;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane
DE protein) (CRP) (60 kDa cysteine-rich OMP).
GN OMCB OR OMP2 OR OMP2B OR CT443.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B/JALI20/OT;
RX MEDLINE=90128208; PubMed=2612891;
RA Watson M.W., Lambden P.R., Ward M.E., Clarke I.N.;
RT "Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein:
RL sequence homology between trachoma and LGV biovars.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B/TW-05/OT;
RX MEDLINE=91141306; PubMed=2287277;
RA Allen J.E., Cerrone M.C., Beatty P.R., Stephens R.S.;
RT "Cysteine-rich outer membrane proteins of Chlamydia trachomatis
RT display compensatory sequence changes between biovariants.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
```

|                       |  |   |                                |             |  |  |
|-----------------------|--|---|--------------------------------|-------------|--|--|
|                       | INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.   |   |                                |             |  |  |
| -                     | SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.   |   |                                |             |  |  |
| This                  | SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensedisb-sib.ch). |   |                                |             |  |  |
| EMBL;                 | X53510;  | CAA37588.1;   | --                             |             |  |  |
| EMBL;                 | AE001317;  | AAC68042.1;   | ALT_INIT.                      |             |  |  |
| PIR;                  | S11673;  | S11673.   | Chlam_OMP6.                    |             |  |  |
| InterPro;             | IPR003506;   |   | Chlam_OMP6.                    |             |  |  |
| PRINTS;               | PR01336;   | CHLAMIDIAOM6.   |                                |             |  |  |
| Outer                 | membrane; TRANSMEMBRANE; Signal; Virulence; Complete proteome.   |   |                                |             |  |  |
| FT                    | SIGNAL   | 1..22   | POTENTIAL.                     |             |  |  |
| ET                    | PROPEP   | 23..40  |                                |             |  |  |
| ET                    | CHAIN  | 41..547   | 60 KDA OUTER MEMBRANE PROTEIN. |             |  |  |
| FT                    | VARIANT  | 233..233  | I -> V (IN STRAIN B/TW-05/OT). |             |  |  |
| SEQUENCE              | 547 AA;  | 58694 MW;   | 42719B4BCEDCCEA CRC64;         |             |  |  |
| Query Match           | 73.2%;   | Score 2160.5;   | DB 1;                          | Length 547; |  |  |
| Best Local Similarity | 71.7%;   | Pred. No. 2.2e-154;   | Indels 15;                     | Gaps 4;     |  |  |
| Matches 401;          | Conservative 75;   | Mismatches 68;  |                                |             |  |  |
| QY                    | 1  | MSKLLRRVVTVLALTSMAFCSPAGGIEAAVASESLITKIIVASAEETKPAPVPMTAKKVRLVR | 60                             |             |  |  |
| Dd                    | 1  | MKKLLRRATVFIAFVAVSLAFSGVLETSMAESLSNTVISLADTKAK--DNTSHSKKAR      | 58                             |             |  |  |
| QY                    | 61   | RN---KQPVEQRSGRAFCDKEFPYCEEGRCQPVGAQQESCGLYSKYVNDCNCNVEIOS      | 117                            |             |  |  |
| Dd                    | 59   | KHNSKETPVDR-----KEVAPYESKA--TGPQODSCFGMYTVKVNDDRRNVLTQA         | 108                            |             |  |  |
| QY                    | 118  | VPEYATVGSYPYTEILAIIGCKDCVDVVIITOQLPCEAEFVSDDPETPTSDGKLWKIDRL    | 177                            |             |  |  |
| Dd                    | 109  | VPEYATVGSYPYEITATGKRDCVDVIIITOQLPCEAEFFVRSDPATTPADGKLWKIDRL     | 168                            |             |  |  |
| QY                    | 178  | GAGDKCKIITVWKPLKEGGCFATAATVCACPSELRSYTKCGOPATCIKOEGPDCACLRCPCV  | 237                            |             |  |  |
| Dd                    | 169  | GQGEKSKITVWKPLKEGGCFATAATVCACPSELRSYTKCGOPATCIKOGPENACLRCPPV    | 228                            |             |  |  |
| QY                    | 238  | YKEVNTGSAIARNVTVDNPDPDGYSHASGORVLSFNLDMPRGDKKKVTFVEFCPORRG      | 297                            |             |  |  |
| Dd                    | 229  | YKININVNOGTARANVVENHPDPGYAHSSGORVLITFLGDMQPGEHRTIVFECPCLKRG     | 288                            |             |  |  |
| QY                    | 298  | QITNVATVTYCGGHKCSANVTTVYNPECQVNIAGDWSYCYCKPYEYSISVSNPGDLVLH     | 357                            |             |  |  |
| Dd                    | 289  | RATNIATVSYCGGHKHKTASVTVINPECQVSIAGDWSYCYCKPYEVYISVSNPGDLVLH     | 348                            |             |  |  |
| QY                    | 358  | DVVTQDTLPSTGVTLVLEAPGGEICCNKVWRIRKEMCPGETLOFKLVKVAQVPGRETNQAV   | 417                            |             |  |  |
| Dd                    | 349  | DVWVEDTLPSGVTVLEAAGAQLSCNKVWTVKELNPGESLYKVLVRAQTPOGFNNVVV       | 408                            |             |  |  |
| QY                    | 418  | TSESCGTCTCAEATTTHHWGLAATHMCVLTDNDPICVGENTVYRICVTNRGSAEDTNVS     | 477                            |             |  |  |
| Dd                    | 409  | KSCDCGTCTCAEATTTWKGVAATHMCVWDTCDPVCGENTVYRICVTNRGSAEDTNVS       | 468                            |             |  |  |
| QY                    | 478  | LILKFSKELOPATSGBPRTGISGNVTVFDPALPKLGSKESVEFSVTLTGIAIPGDARGEAI   | 537                            |             |  |  |
| Dd                    | 469  | LMKFPSKELOQVFSFGTKGTITGNTVVFDSLPLRLGSKETVEFSVTLKAVSAGARGEAI     | 528                            |             |  |  |
| QY                    | 538  | LSSDILTSPVSDTENTHYV   | 556                            |             |  |  |
| Dd                    | 529  | LSSDILTVPVSDTENTHIY   | 547                            |             |  |  |
| RESULT                | 6  |   |                                |             |  |  |
| OM6L_CHLTR            |  |   |                                |             |  |  |
| TO_OM6L_CHLTR         |  |   |                                |             |  |  |
| STANDARD:             |  |   |                                |             |  |  |
| PRT:                  | 547 AA.  |   |                                |             |  |  |

P21354; P18596;  
01-NOV-1990 (Rel. 16, Created)  
01-MAY-1991 (Rel. 18, Last sequence update)  
30-MAY-2000 (Rel. 39, Last annotation update)  
DE 60 kDa outer membrane protein, serovars L1/L2/L3 precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).  
GN OMP2 OR OMP2B.  
DN Chlamydia trachomatis.  
OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OC NCBI\_TaxID=813;  
OX [1]  
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RP STRAIN=L2/434BU;  
RC MEDLINE=89123030; PubMed=2914847;  
RX Allen J.E., Stephens R.S.;  
RA "Identification by sequence analysis of two-site posttranslational processing of the cysteine-rich outer membrane protein 2 of Chlamydia trachomatis serovar L2.";  
RT J. Bacteriol. 171:285-291(1989).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=404 / SEROVAR L3;  
RC MEDLINE=G1147205; PubMed=1997423;  
RX de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;  
RA "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";  
RT Infect. Immun. 59:1196-1201(1991).  
RL [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=L1/440LIN;  
RC MEDLINE=89138006; PubMed=3066701;  
RX Clarke I.N., Ward M.E., Lambden P.R.;  
RA "Molecular cloning and sequence analysis of a developmentally regulated cysteine-rich outer membrane protein from Chlamydia trachomatis";  
RT Gene 71:307-314(1988).  
RL -! FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBs OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.  
CC -! SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenseseis@sib.ch).

-----

EMBL: X23001; AAA23152.1; -  
DR EMBL: MS4390; CAA38261.1; -  
DR EMBL: M35148; AAA23119.1; ALT\_INIT.  
PIR: A32244; A32244.  
DR PIR: S18981; S18981.  
DR PIR: A43584; A43584.  
DR PIR: JT0419; JT0419.  
DR Siena-2DPAGE; P21354; -  
DR InterPro: IPR003506; Chlam\_OMP6.  
DR PRINTS: PR01336; CHLAMIDIAOM6.  
KW Outer membrane; Transmembrane; Signal; Virulence.  
FT SIGNAL 1 22  
FT PROPEP 23 40  
FT CHAIN 41 547  
FO SEQUENCE 547 AA; 58782 MW; 78CBE41CCF98472D CRC64;

|                       |                  |                     |            |             |
|-----------------------|------------------|---------------------|------------|-------------|
| Query Match           | 72.2%;           | Score 2129.5;       | DB 1;      | Length 547; |
| Best Local Similarity | 70.7%;           | pred. No. 4.5e-152; |            |             |
| Matches 395;          | Conservative 77; | Mismatches 72;      | Indels 15; | Gaps 4;     |

```

QY 1 MSLKIRRVTVTLTSMASCFASGIEAAVESLITKIVASAEKTPAPVPMTAKKRVLR 60
Db 1 MNKLRRAVTVFAVTSVSLFASGVLSTMAEFSTNVISLADTKAK--DNTSHKSKAR 58
QY 61 RN---KOPVEOKSGAFCDKEFYCECEGRQCPVEAQSCYGRLYSVKVNDDCNVEIQOS 117
Db 59 KNSKSTPVNR-----KKVAPVHESKA--TGPQKDSFCFGRMYTVKVNDDRVNVEITQA 108
QY 118 VPEYATGSPYPIELAIKGRKDCVDVLTQOLPCEAEFVSSDPPTTSDGKLWVKIDRL 177
Db 109 VPKYATGSPVPVEITATGRKDCVDVLTQOLPCEAEFVSSDPPTTADGKLWVKIDRL 168
QY 178 GAGDKCITVWVPLKCGCCFTATVYACACPELRSYTCGQPAICIKQEGPDCACLRCPVC 237
Db 169 GQSEKSKITVWVPLKCGCCFTATVYACACPEIRSVTCGQPAICVKQEGFENACLRCPVV 228
QY 238 YKIEVNTGSAIARNVTDNPNPDGYSHASSORVLSFNLDGMRPGDKKVTVEFCPPORR 297
Db 229 YKINVNOGTATARNVVENPVPDSVAHSGQRVLFTLGDMPGEGHRTITVEFCPLKRG 288
QY 298 QITNVATVYCGGHKCSANVTTVVNEPCVQVNIAGDWSYVCKPVEYSVSNPGDLVLH 357
Db 289 RATNIAMVYCGGHKCSANVTTVVNEPCVQVNIAGDWSYVCKPVEYSVSNPGDLVLH 348
QY 358 DVVIOTPLSGVTVLEAPGGETCCNKVVRKEMCPGETLQKLVKYAQVGRFTNQAV 417
Db 349 DVVVKDTLSPGTVVLEAAGAQISCNKVVTVVRELNPGESLQYKVLVRAQTPGQFTNNVV 408
QY 418 TSESNGCTSCAETTHWKGLAATHMVCVLDTNDPICVGVNTVYICVTRNRSADETNVS 477
Db 409 KSCSDGCTSCAETTHWKGLAATHMVCVLDTNDPICVGVNTVYICVTRNRSADETNVS 468
QY 478 LILKESKELQPIASSGPKTGTSIGNTVVFDALPKLGSKESEFVSTLKGIPGCDARGAI 537
Db 469 LMLKESKELQPVPSGPKTGITGTVVFDLSPLRGSKETVEFVSTLKAVSAGDARGAI 528
QY 538 LSSDPLTSPVSDTENTHYI 556
Db 529 LSSDPLTSPVSDTENTHYI 547

RESULT 7
SERR_DROME
ID SERR_DROME STANDARD; PRT; 1408 AA.
AC P18168;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Serrate protein precursor (Beaded protein).
GN SER OR BD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=91347903; PubMed=1840519;
RA Thomas U., Speicher S.A., Knust E.;
RT "The Drosophila gene serrate encodes an EGF-like transmembrane
RT protein with a complex expression pattern in embryos and wing
RT discs.";
RL Development 111:749-761(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9109666; PubMed=2125287;
RA Fleming R.J., Scottage T.N., Diederich R.J., Artavanis-Tsakonas S.;
RT "The gene Serrate encodes a putative EGF-like transmembrane protein
RT essential for proper ectodermal development in Drosophila
RT melanogaster.";
RL Genes Dev. 4:2188-2201(1990).

```

```

CC -!- FUNCTION: ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT. SERRATE
CC MAY REPRESENT AN ELEMENT IN A NETWORK OF INTERACTING MOLECULES
CC OPERATING AT THE CELL SURFACE DURING THE DIFFERENTIATION OF
CC CERTAIN TISSUES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: APPEARS TO BE RESTRICTED EXCLUSIVELY TO
CC CELLS OF ECTODERMAL ORIGIN.
CC -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
CC NEUROGENIC GENES.
CC -!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN
CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
CC COMPLETE FOR BINDING WITH THE NOTCH PROTEIN.
CC -!- SIMILARITY: CONTAINS 14 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X56811; CAA40148.1; -
CC DR EMBL: M35759; AAA28938.1; -
CC DR PIR: A36666; A36666.
CC DR PIR: S16878; S16878.
CC DR HSSP: P00743; ICCF.
CC DR FlyBase: FBgn004197; Ser.
CC DR InterPro: IPR000152; Asx_hydroxyl.
CC DR InterPro: IPR001774; DSL.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR000742; EGF-2.
CC DR InterPro: IPR001881; EGF-Ca.
CC DR InterPro: IPR001438; EGF-II.
CC DR Pfam: PF01414; DSL; 1.
CC DR Pfam: PF00008; EGF; 11.
CC DR PRINTS: PR00010; EGFBL00D.
CC DR SMART: SM00051; DSL; 1.
CC DR SMART: SM00179; EGF_CA; 7.
CC DR SMART: SM00001; EGF_like; 5.
CC DR PROSITE: PS00010; ASX_HYDROXYL; 7.
CC DR PROSITE: PS00022; EGF_1; 14.
CC DR PROSITE: PS01186; EGF_2; 8.
CC DR PROSITE: PS01187; EGF_CA; 5.
CC KW Differentiation; Repeat; EGF-like domain; Transmembrane;
CC Glycoprotein; Signal.
CC FT SIGNAL 1 83
CC FT CHAIN 84 1408
CC FT DOMAIN 84 1223
CC FT TRANSMEM 1224 1249
CC FT DOMAIN 1250 1408
CC FT DOMAIN 284 317
CC FT DOMAIN 315 349
CC FT DOMAIN 351 389
CC FT DOMAIN 391 489
CC FT DOMAIN 407 476
CC FT DOMAIN 491 527
CC FT DOMAIN 529 609
CC FT DOMAIN 611 646
CC FT DOMAIN 648 684
CC FT DOMAIN 686 721
CC FT DOMAIN 723 737
CC FT DOMAIN 737 769
CC FT DOMAIN 799 835
CC FT DOMAIN 837 877
CC FT DOMAIN 879 915
CC FT DOMAIN 917 953
CC FT DISULFID 288 299
CC FT DISULFID 292 305
CC FT DISULFID 307 316

```



|    |      |   |      |
|----|------|---|------|
| Db | 679  | SGSLCEEAKENCTPSPCLGHCINTPEGYCHCPDPDRAGKHCEQURPLCSQPPCNGCEFA | 738  |
| QY | 222  | -----IKQGPDC-----ACLRCPV---CYKIEVYVNTGSA                    | 250  |
| Db | 739  | NVSLATSATTTTTTTTATTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHTGTGTC  | 797  |
| QY | 251  | RNYTVDNVPV-----DG-----YSHASQRYLS-----FNLGDMRPG--            | 282  |
| Db | 798  | HNLNECSPNCRNGGICLDGDFTECEMSGWTGKRCSEATGCGYAGOCQNGGTCHPGAP   | 857  |
| QY | 283  | DKV-----FTVEFCPQRRGQTNVATVYCGGHKCSANVTVWNEPCVQVNISGA        | 333  |
| Db | 858  | DKALQPHCRCAPGWTGLFCAE-----AIDCRQCPCHNGGT-----CE-----SGA     | 898  |
| QY | 334  | DW-SYVCKPVEYTSVSNPGLVLHDVVIODTLPSTVTVLEAPGGEITCCNVVRIKEMC   | 392  |
| Db | 899  | GWFRVCV-----AQGSGPDCRINVNCSPOPCGGATCIDGGYSC-----IC          | 943  |
| QY | 393  | PGETLQFLVKAQVPGRETNQAVTSSESNGTCTCAET-----TTHWKGLAATH        | 443  |
| Db | 944  | P-----PGRHLRCELLISDPKSAQONASNTISPYTALNRSONWLDIALTG          | 989  |
| QY | 444  | MCVLDTNPICGVGVYRICVTNRGSAEDTNVSLILKFSKLOPIASSGPTKGTISGNT    | 503  |
| Db | 990  | RTEDDENCNACVEN-----GTSRCTNMLCGLPNCYKVDPLSKSSNLGVCVKOHE      | 1039 |
| QY | 504  | VVFDAL 509  |      |
| Db | 1040 | VCVPAL 1045   |      |

RESULT 8

ANNU\_SCHAM

ID

ANNU\_SCHAM

STANDARD;

PRT;

772 AA.

AC

P51-021-1996 (Rel. 34, Created)

01-OCT-1996 (Rel. 34, Last annotation update)

DT

15-JUL-1999 (Rel. 38, Last annotation update)

DT

Annulin (protein-glutamine gamma-glutamyltransferase) (EC 2.3.2.13)

DE

(Transglutaminase).

DE

Schistocerca americana (American grasshopper).

OS

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC

Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;

OC

Acridomorpha; Acridoidea; Acrididae; Schistocerca.

OX

NCBI\_Taxid=7009;

[1]

SEQUENCE FROM N.A.

RP

MEDLINE=93050772; PubMed=1358727;

RX

Singer M.A., Hortsch M., Goodman C.S., Bentley D.;

RA

'Annulin, a protein expressed at limb segment boundaries in the

RT

grasshopper embryo, is homologous to protein cross-linking

RT

transglutaminases.';

RL

Dev. Biol. 154:143-159(1992).

CC

!- FUNCTION: PARTICIPATES IN MORPHOGENETIC ACTIVITIES OF THE CELLS,

CC

MAYBE BY STABILIZING THE MEMBRANE OR SUBCORTICAL STRUCTURES OF

CC

CELLS THAT ARE UNDER MECHANICAL STRESS. PROBABLY CATALYZES THE

CC

CROSS-LINKING OF PROTEINS AND THE CONJUGATION OF POLYAMINES TO

CC

PROTEINS.

CC

!- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N5-

CC

alkylglutamine + NH(3).

CC

!- COFACTOR: CALCIUM IS REQUIRED FOR TRANSGLUTAMINASE ACTIVITY (BY

CC

SIMILARITY).

CC

!- SUBCELLULAR LOCATION: INTRACELLULAR AND PERIPHERALLY ASSOCIATED

CC

WITH THE INNER LEAFLET OF THE CELL MEMBRANE, USING A FATTY ACID

CC

LINKAGE.

CC

!- TISSUE SPECIFICITY: HAS AN ANNULAR, OR RING-LIKE EXPRESSION

CC

PATTERN IN EPITHELIAL ANNULI OF DEVELOPING LIMB SEGMENT BOUNDARY

CC

CELLS. IN EMBRYOS, IT IS SEEN IN GASTRULATING CELLS, IN CELLS

CC

SURROUNDING RAPIDLY DIVIDING NEUROBLASTS, AND IN MUSCLE PIONEER

CC

CELLS INVAGINATING TO FORM APODEMES.

CC

!- DEVELOPMENTAL STAGE: EXPRESSION OF THIS PROTEIN IN EMBRYOS AND

CC

LIMBS IS ASSOCIATED WITH AREAS UNDERGOING MOVEMENTS, MORPHOGENETIC

CC

REARRANGEMENTS, OR RAPID CELL DIVISION. EXPRESSION OF ANNULIN

CC PRECEDES THE FIRST MORPHOLOGICAL SIGNS OF SEGMENTATION IN THE  
CC DEVELOPING LIMBS.  
CC  
CC -|- SIMILARITY BELONGS TO THE TRANSGLUTAMINASE FAMILY.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

|    |  |
|----|--|
| CC | EMBL; M92291; AAA29806.1; --                       |
| DR | HSSP; P00488; 1GGU.                                |
| DR | InterPro; IPR002931; Transglut_core.               |
| DR | InterPro; IPR001102; Transglutmunse.               |
| DR | Pfam; PF00927; Transglutamin_C; 1.                 |
| DR | Pfam; PF00868; Transglutamin_N; 1.                 |
| DR | Pfam; PF01841; Transglut_core; 1.                  |
| DR | SMART; SM00460; TGG; 1.                            |
| DR | PROSITE; PS00547; TRANSGLUTAMINASES; 1.            |
| KW | Transferase; Acyltransferase; Calcium-binding.     |
| FT | LIPID 4 5  |
| FT | POTENTIAL.   |
| FT | POTENTIAL.   |
| FT | ACT SITE 341 341                                   |
| FT | DOMAIN 567 570                                     |
| FT | BY SIMILARITY.                                     |
| FT | POLY-VAL.  |
| SQ | SEQUENCE 772 AA; 85941 MW; FASA3CE6A7C4E394 CRC64; |

|                       |        |   |   |                     |
|-----------------------|--------|---|---|---------------------|
| Query Match           | 4.4%;  | Score 129;  | DB 1;   | Length 772;         |
| Best Local Similarity | 23.7%; | Pred. No. 0.036;  |   |                     |
| Matches               | 79;    | Conservative 55;  | Mismatches 118;                                 | Indels 82; Gaps 19; |
| QY                    | 41     | SAETKPAPWMTAKKVLVRNK----  | QPEQKSQSGAFCDKEF-----                           | YP--CEGRGQCP 89     |
| Db                    | 476    | SGTQP-----  | LKLTKDMLGIGNISTKAVGRFQREDITNTYKYPEKSVEERAAM 526 |                     |
| QY                    | 90     | VEA--QOESCYGRLYSVKYNDDCN--VEIQSQSPEYATGSPYPIELTAGKDKCDVDVIT 146 |   |                     |
| Db                    | 527    | LKALROSESLSRYV--LNEDEFNHFNFELRDDIVIGSPFVVVWKNRSNQODYT 583       |   |                     |
| QY                    | 147    | QQLPCAEAFVSSDPETPTSDOKLWVKIDRL---                               | GAGDKCKITV---WVKPLKEGCGFT 199                   |                     |
| Db                    | 584    | VLLRVDTVLY-----   | TGHVKDGVKKKEVERLKAGAVEIRDVSVDYKYHLVDQCAFN 638   |                     |
| QY                    | 200    | AATVCACPELRSYTKGQ-----  | PAICIKOEGPDCACLRCPVCYKIEVNTGSGAIARN 252         |                     |
| Db                    | 639    | IACLATVHD--TWVEYFAQDDFVRVPDTKIKUEG-----                         | PV-----QGQMSAV 682                              |                     |
| QY                    | 253    | VTVNDNVP-----   | DGYSHASQQRV--LSFNLGDMRPGDKKVFTEVFCPPQRGQITN 301 |                     |
| Db                    | 683    | ATLKNPLPTPVKKGQFLIEBPGGIAGTQKIKLSQNTA---                        | PGEASVNFKFPKPYDGR--- 736                        |                     |
| QY                    | 302    | VATVTYCGGHKCSANVTTVN---   | EPCVOVNISG 332                                  |                     |
| Db                    | 737    | -ATIAAKSSKELDDVDGFLNFWPEPKKEVANGTG 769                          |   |                     |

|   |                                   |
|---|-----------------------------------|
| RESULT  | 9                                 |
| ID_CRB_DROME  |                                   |
| ID_CRB_DROME  | PRT; 2139 AA.                     |
| AC_P10040;  | STANDARD;                         |
| DT_01-MAR-1989  | (Rel. 10, Created)                |
| DDT_01-MAY-1991   | (Rel. 18, Last sequence update)   |
| DDT_01-MAR-2002   | (Rel. 41, Last annotation update) |
| DEDE  | Crumbs protein precursor (95F).   |
| CGN_CRB.  |                                   |
| CCN_Drosophila melanogaster   | (Fruit fly).                      |
| OOC_Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;         |                                   |
| OOC_Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; |                                   |
| OOC_Ephydroidea; Drosophilidae; Drosophila.                               |                                   |
| OOC_NCBITaxID=7227;   |                                   |
| RNRN  | [1]                               |

SEQUENCE FROM N.A.  
RC STRAIN-OREGON-R; TISSUE=Embryo;  
RX MEDLINE=90263104; PubMed=2344615;  
RA Tepass U., Theres C., Knust E.;  
RT "Crumbs encodes an EGF-like protein expressed on apical membranes of  
RT drosophila epithelial cells and required for organization of  
RL epithelia";  
RL Cell 61:787-799(1990).  
[2]  
RP SEQUENCE OF 1663-1955 FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=87218537; PubMed=3107986;  
RA Knust E., Dietrich U., Tepass U., Bremer K.A., Welgel D.,  
RA Vaessin H., Campos-Ortega J.A.;  
RT "EGF homologous sequences encoded in the genome of Drosophila  
RL melanogaster, and their relation to neurogenic genes.";  
RL EMBO J. 6:761-766(1987).  
CC -!- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA,  
CC POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL  
CC POLARITY. IT MAY ACT AS A SIGNAL.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).  
CC -!- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M33753; AAA28428.1; ALT\_SEQ.  
DR EMBL: X05144; CAA38793.1; -  
DR PIR: B26637; B26637.  
DR PIR: A35672; A35672.  
DR HSP: P00740; IIXA.  
DR FlyBase: Fgn0000368; crb.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF-2.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_II.  
DR InterPro: IPR001791; Laminin\_G.  
DR Pfam: PF00008; EGF; 26.  
DR Pfam: PF00054; laminin\_G; 3.  
DR PRINTS: PR00010; EGFBL00D.  
DR SMART: SM00179; EGF\_CA; 11.  
DR SMART: SM00001; EGF\_like; 16.  
DR SMART: SM00282; LamG; 3.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 15.  
DR PROSITE: PS00022; EGF\_1; 26.  
DR PROSITE: PS01186; EGF\_2; 17.  
DR PROSITE: PS01187; EGF\_CA; 15.  
DR PROSITE: PS50025; LAM\_G\_DOMAIN; 3.  
KW Differentiation; Repeat; EGF-like domain; Transmembrane;  
KW Glycoprotein; Signal; Phosphorylation.  
FT SIGNAL 1 90  
FT CHAIN 91 2139  
FT DOMAIN 91 2084  
FT TRANSMEM 2085 2111  
FT DOMAIN 2112 2139  
FT DOMAIN 267 303  
FT DOMAIN 306 343  
FT DOMAIN 348 386  
FT DOMAIN 388 425  
FT DOMAIN 427 463  
FT DOMAIN 464 500  
FT DOMAIN 501 532  
FT DOMAIN 545 581  
FT DOMAIN 582 611  
FT DOMAIN 609 646  
FT CRUMBS PROTEIN,  
FT EXTRACELLULAR (POTENTIAL).  
FT POTENTIAL.  
FT CYTOPLASMIC (POTENTIAL).  
FT EGF-LIKE 1.  
FT EGF-LIKE 2.  
FT EGF-LIKE 3.  
FT EGF-LIKE 4.  
FT EGF-LIKE 5.  
FT EGF-LIKE 6.  
FT EGF-LIKE 7.  
FT EGF-LIKE 8.  
FT EGF-LIKE 9.  
FT EGF-LIKE 10.  
FT CALCIUM-BINDING (POTENTIAL).



```
RESULT 10
GLP1_CAEEL STANDARD; PRT; 1295 AA.
AC PI3508;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE GLP-1 protein precursor.
GN GLP-1 OR EMB-33 OR F02A9.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=89336787; PubMed=2758466;
RA Yochem J., Greenwald I.;
RT "glp-1 and lin-12, genes implicated in distinct cell-cell
RL interactions in C. elegans, encode similar transmembrane proteins.";
RN Cell 58:553-563(1989).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lighning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans";
RL Nature 368:32-38(1994).
RN [3]
RP DELETION OF 1174-1295.
RX MEDLINE=91351288; PubMed=1881436;
RA Mango S.E., Maine E.M., Kimble J.;
RT "Carboxy-terminal truncation activates glp-1 protein to specify
RL vulval fates in Caenorhabditis elegans.";
RN Nature 352:811-815(1991).
RP [4]
RP CHARACTERIZATION OF FUNCTION OF THE ANK-REPEATS.
RX MEDLINE=93354444; PubMed=8350921;
RA Roehl H., Kimble J.;
RT "Control of cell fate in C. elegans by a GLP-1 peptide consisting
RL primarily of ankyrin repeats.";
RN Nature 364:632-635(1993).
RP [5]
RP FUNCTION.
RX MEDLINE=94208066; PubMed=8156602;
RA Mello C.C., Draper B.W., Priess J.R.;
RT "The maternal genes apx-1 and glp-1 and establishment of
RL dorsal-ventral polarity in the early C. elegans embryo.";
RN Cell 77:95-106(1994).
RP [6]
RP FUNCTION: INVOLVED IN THE SPECIFICATION OF THE CELL FATES OF THE
CC BLASTOMERES, ABA AND APA. PROPER SIGNALING BY GLP-1 INDUCES ABA
CC DESCENDANTS TO PRODUCE ANTERIOR PHARYNGEAL CELLS, AND APA
CC DESCENDANTS TO ADOPT A DIFFERENT FATE. CONTRIBUTES TO THE
CC ESTABLISHMENT THE DORSAL-VENTRAL AXIS IN EARLY EMBRYOS.
CC [7]
RP SUBCELLULAR LOCATION: Type I membrane protein.
CC [8]
RP DEVELOPMENTAL STAGE: ACTS ON ABA DEVELOPMENT DURING 4-CELL AND
CC 12-CELL STAGES, AND ON ABA DEVELOPMENT DURING 12-CELL AND 28-CELL
CC STAGES.
CC [9]
RP SIMILARITY: HIGH, TO C.ELEGANS LIN-12.
CC [10]
RP SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
CC [11]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [12]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [13]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [14]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [15]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [16]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [17]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [18]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [19]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [20]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [21]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [22]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [23]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [24]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [25]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [26]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [27]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [28]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [29]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [30]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [31]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [32]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [33]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [34]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [35]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [36]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [37]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [38]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [39]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [40]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [41]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [42]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [43]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [44]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [45]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [46]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [47]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [48]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [49]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [50]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [51]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [52]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [53]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [54]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [55]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [56]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [57]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [58]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [59]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [60]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [61]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [62]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [63]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [64]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [65]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [66]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [67]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [68]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [69]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [70]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [71]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [72]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [73]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [74]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [75]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [76]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [77]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [78]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [79]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [80]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [81]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [82]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [83]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [84]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [85]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [86]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [87]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [88]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [89]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [90]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [91]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [92]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [93]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [94]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [95]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [96]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [97]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [98]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
CC [99]
RP SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC [100]
RP SIMILARITY: CONTAINS 5 ANK REPEATS.
```

|    |  |   |     |                                     |   |
|----|--|---|-----|-------------------------------------|---|
| FT | DISULFID                               | 194   | 206 | BY SIMILARITY.                      | TRBM_MOUSE  |
| FT | DISULFID                               | 201   | 218 | BY SIMILARITY.                      | ID TRBM_MOUSE   |
| FT | DISULFID                               | 220   | 229 | BY SIMILARITY.                      | AC P15306;  |
| FT | DISULFID                               | 236   | 248 | BY SIMILARITY.                      | 01-APR-1990 (Rel. 14, Created)  |
| FT | DISULFID                               | 242   | 257 | BY SIMILARITY.                      | 01-APR-1990 (Rel. 14, Last sequence update)                                   |
| FT | DISULFID                               | 259   | 268 | BY SIMILARITY.                      | 16-OCT-2001 (Rel. 40, Last annotation update)                                 |
| FT | DISULFID                               | 275   | 286 | BY SIMILARITY.                      | Thrombomodulin precursor (Fetomodulin) (TM).                                  |
| FT | DISULFID                               | 280   | 296 | BY SIMILARITY.                      | THBD.   |
| FT | DISULFID                               | 298   | 307 | BY SIMILARITY.                      | GN Mus musculus (Mouse).  |
| FT | DISULFID                               | 373   | 384 | BY SIMILARITY.                      | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |
| FT | DISULFID                               | 378   | 394 | BY SIMILARITY.                      | OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.         |
| FT | DISULFID                               | 396   | 405 | BY SIMILARITY.                      | OX NCBI_TaxID=10090;  |
| FT | DISULFID                               | 411   | 422 | BY SIMILARITY.                      | RP SEQUENCE FROM N.A. PubMed-2536925;   |
| FT | DISULFID                               | 416   | 431 | BY SIMILARITY.                      | RA Dittman W.A., Majerus P.W.;  |
| FT | DISULFID                               | 433   | 442 | BY SIMILARITY.                      | RT "Sequence of a cDNA for mouse thrombomodulin and comparison of the         |
| FT | DISULFID                               | 450   | 461 | BY SIMILARITY.                      | RT predicted mouse and human amino acid sequences.";                          |
| FT | DISULFID                               | 455   | 467 | BY SIMILARITY.                      | RL Nucleic Acids Res. 17:802-802(1989).                                       |
| FT | DISULFID                               | 469   | 478 | BY SIMILARITY.                      | RN N-LINKED (GLCNAC. . .) (POTENTIAL).  |
| FT | CARBOHYD                               | 244   | 244 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RP MEDLINE-89008498; PubMed-2844823;  |
| FT | CARBOHYD                               | 245   | 245 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RA Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;                        |
| FT | CARBOHYD                               | 333   | 333 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RT "The structure and function of mouse thrombomodulin. Phorbol               |
| FT | CARBOHYD                               | 381   | 381 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RT myristate acetate stimulates degradation and synthesis of                  |
| FT | CARBOHYD                               | 609   | 609 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RT thrombomodulin without affecting mRNA levels in hemangioma cells.";        |
| FT | CARBOHYD                               | 675   | 675 | N-LINKED (GLCNAC. . .) (POTENTIAL). | RL J. Biol. Chem. 263:15815-15822(1988).                                      |
| SQ | SEQUENCE                               | 1295 AA; 144078 MW; 422AAD0A2DEEF3B4 CRC64; |     |                                     | CC -!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR       |
|    |  |   |     |                                     | CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS                |
|    |  |   |     |                                     | CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE              |
|    |  |   |     |                                     | CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA                 |
|    |  |   |     |                                     | CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,             |
|    |  |   |     |                                     | CC FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF              |
|    |  |   |     |                                     | CC THROMBIN GENERATED.  |
|    |  |   |     |                                     | CC -!- SUBCELLULAR LOCATION: Type I membrane protein.                         |
|    |  |   |     |                                     | CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING       |
|    |  |   |     |                                     | CC THROMBOMODULIN (BY SIMILARITY).  |
|    |  |   |     |                                     | CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.                               |
|    |  |   |     |                                     | CC This SWISS-PROT entry is copyright. It is produced through a collaboration |
|    |  |   |     |                                     | CC between the Swiss Institute of Bioinformatics and the EMBL outstation -    |
|    |  |   |     |                                     | CC the European Bioinformatics Institute. There are no restrictions on its    |
|    |  |   |     |                                     | CC use by non-profit institutions as long as its content is in no way         |
|    |  |   |     |                                     | CC modified and this statement is not removed. Usage by and for commercial    |
|    |  |   |     |                                     | CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ |
|    |  |   |     |                                     | CC or send an email to license@isb-sib.ch).                                   |
|    |  |   |     |                                     | CC -----  |
|    |  |   |     |                                     | DR EMBL: X14432; CAA32597.1; -  |
|    |  |   |     |                                     | DR PIR: A32001; A32001.   |
|    |  |   |     |                                     | DR PIR: S08488; S08488.   |
|    |  |   |     |                                     | DR HSSP: P07204; LEGT.  |
|    |  |   |     |                                     | DR MGD: MGI:98736; Thbd.  |
|    |  |   |     |                                     | DR InterPro: IPR000152; Asx_hydroxyl.   |
|    |  |   |     |                                     | DR InterPro: IPR000561; EGF-like.   |
|    |  |   |     |                                     | DR InterPro: IPR001881; EGF_Ca.   |
|    |  |   |     |                                     | DR InterPro: IPR001491; Thrombomodulin.                                       |
|    |  |   |     |                                     | DR InterPro: IPR001304; lectin_c.   |
|    |  |   |     |                                     | DR Pfam: PF00008; EGF; 4.   |
|    |  |   |     |                                     | DR Pfam: PF00059; lectin_c; 1.  |
|    |  |   |     |                                     | DR PRINTS: PR00907; THROMBOMODULN.  |
|    |  |   |     |                                     | DR SMART: SM00034; EGF_Ca; 1.   |
|    |  |   |     |                                     | DR SMART: SM00179; EGF_Ca; 1.   |
|    |  |   |     |                                     | DR SMART: SM00001; EGF_Like; 5.   |
|    |  |   |     |                                     | DR PROSITE: PS00010; ASX_HYDROXYL; 2.   |
|    |  |   |     |                                     | DR PROSITE: PS00022; EGF_1; FALSE_NEG.  |
|    |  |   |     |                                     | DR PROSITE: PS01186; EGF_2; 3.  |
|    |  |   |     |                                     | DR PROSITE: PS01187; EGF_Ca; 2.   |
|    |  |   |     |                                     | DR PROSITE: PS50041; C-TYPE LECTIN 2; 1.                                      |
|    |  |   |     |                                     | DR Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;      |
| KW | Glycoprotein; Signal; EGF-like domain. |   |     |                                     |   |
| FT | SIGNAL                                 | 1   | 16  | POTENTIAL.                          |   |
| FT | CHAIN                                  | 17  | 577 | THROMBOMODULIN.                     |   |
| FT | DOMAIN                                 | 17  | 517 | EXTRACELLULAR (POTENTIAL).          |   |

TRANSMEM 518 541 POTENTIAL.  
FT DOMAIN 542 577 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 543 578 EGF-LIKE 1.  
FT DOMAIN 544 579 EGF-LIKE 2.  
FT DOMAIN 545 580 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 546 581 EGF-LIKE 4.  
FT DOMAIN 547 582 EGF-LIKE 5.  
FT DOMAIN 548 583 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
FT DISULFID 244 255 BY SIMILARITY.  
FT DISULFID 251 264 BY SIMILARITY.  
FT DISULFID 266 279 BY SIMILARITY.  
FT DISULFID 287 295 BY SIMILARITY.  
FT DISULFID 291 307 BY SIMILARITY.  
FT DISULFID 309 322 BY SIMILARITY.  
FT DISULFID 328 339 BY SIMILARITY.  
FT DISULFID 335 348 BY SIMILARITY.  
FT DISULFID 350 361 BY SIMILARITY.  
FT DISULFID 368 377 BY SIMILARITY.  
FT DISULFID 373 387 BY SIMILARITY.  
FT DISULFID 389 403 BY SIMILARITY.  
FT DISULFID 407 416 BY SIMILARITY.  
FT DISULFID 412 424 BY SIMILARITY.  
FT DISULFID 426 438 BY SIMILARITY.  
FT DISULFID 444 454 BY SIMILARITY.  
FT DISULFID 449 463 BY SIMILARITY.  
FT DISULFID 465 479 BY SIMILARITY.  
FT CARBOHYD 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 408 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;

Query Match 3.9%; Score 115.5; DB 1; Length 577;  
Best Local Similarity 21.3%; Pred. No. 0.26; Indels 185; Gaps 32;  
Matches 109; Conservative 55; Mismatches 162; Indels 185; Gaps 32;  
QY 38 IVASAEKPKAP-VPMTAKKVLVRNKPQVQKSRGAFCDKEFY-----PCEGRCPQV 90  
DB 136 VTVSTATEAAPGEPAMEEK-----PCETETQGLFC--EFYFASCRPLTVNTRDPE 184  
QY 91 EAQEQSCYGRLYSVKVNDCNVEICQSVPEYATVSGPYPIELAIKGDVVIQQLP 150  
DB 185 AAHSISTYNTFPVGSAD-----FQLP--VGSSAAVEPLGL-----ELV 222  
QY 151 CEAEFVSSDPETPTSDGKLWKIDRLGAGDKCKITVWVKPLKEGCC-----FTAATVC 204  
DB 223 CRA-----PPGTSEGHAW--EATGA-----WNCVSENGGCEYLCNRSTNEPRC 264  
QY 205 ACPELRSYTKGQPAICIKQEGPCACLRCPVCYKIEVWVTGSAIARNVTVDN--PVDPGY 263  
DB 265 LCPR-----DMDLQADGRSCA-----RPVQVSCNLCBHFVCSNAEVPGSY 305  
QY 264 S-----HASGQVLSFNGLDMRPGDKKVFTEPCPQRRGQITNVATVTCYGGHKCS 314  
DB 306 SCMCETGYOLAADGHR--CEDVDCKQGNP-----CPQ-----LCVNTKGGECEP 349  
QY 315 A-NVTTVNEPCQV--NTSGADWSYCKPV---EYSISVSNPG-----DLVLH 357  
DB 350 CYDGYELVDGECVELLDPCFGSNCFCQCPQVSPDTR--CICAPGFAPKDPHPKCFN 408  
QY 358 DVVI-ODTLPSPGVTVLEAP-----GGEICCNKVVNRKEMCPGETLQKLVVKAQVPR 410  
DB 409 EFTSCPADCPNPTVCEPGEFFILDGSGVCTD-----IDCSGCECTSEC---RNPFGS 460  
QY 411 FTRQAVTSENGCTSCAETTHWKGLAATHMCVLDITNDPICVGTGYVIRCVTRNGS 470  
DB 461 Y-----ECICGPDPTALAGO-----ISKDCDPIPVREDT-----KEEGS 494  
QY 471 AEDTNNVSLLIKFSKELQPTASSGPTKGTISG 501

DB 495 GE-----PPVS---PTPGSPTG 508  
RESULT 12  
ID FB2\_HUMAN STANDARD; PRT; 2911 AA.  
AC P35556;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibrillin 2 precursor.  
GN FB2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94165150; PubMed=8120105;  
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguinetti C.,  
RA Bonadio J., Meham R.P., Ramirez F.,  
RT "Structure and expression of fibrillin-2, a novel microfibrillar  
component preferentially located in elastic matrices.";  
RL J. Cell Biol. 124:855-863(1994).  
RN [2]  
RP SEQUENCE OF 752-1505 FROM N.A.  
RX MEDLINE=91304567; PubMed=1852206;  
RA Lee B., Godfrey M., Vitale E., Hori H., Mattel M.-G., Sarfarazi M.,  
RA Tsipouras P., Ramirez F., Hollister D.,  
RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
two different fibrillin genes.";  
RL Nature 352:330-334(1991).  
RN [3]  
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.  
RX MEDLINE=96083599; PubMed=7493032;  
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;  
RT "Fibrillin-2 (FB2) mutations result in the Marfan-like disorder,  
congenital contractual arachnodactyly.";  
RL Nat. Genet. 11:456-458(1995).  
RN [4]  
RP VARIANTS CCA HIS-1114.  
RX MEDLINE=98407789; PubMed=9737771;  
RA Babcock D., Gasner C., Francke U., Maslen C.;  
RT "A single mutation that results in an asp-to-his substitution and  
partial exon skipping in a family with congenital contractual  
arachnodactyly.";  
RL Hum. Genet. 103:22-28(1998).  
RN [5]  
RP VARIANTS CCA PHE-1141 AND TRP-1252.  
RX MEDLINE=20259236; PubMed=10797416;  
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,  
RA Godfrey M.;  
RT "Two novel fibrillin-2 mutations in congenital contractual  
arachnodactyly.";  
RL Am. J. Med. Genet. 92:7-12(2000).  
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE  
THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.  
CC -!- DISEASE: DEFECTS IN FB2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL  
ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS  
PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE  
AORTA AND THE EYES.  
CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING  
EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```
Db 317 DI---DECSIIPIGICETGECSTNTVGSYFCVCP--RGVVTSTGSRCIDQRTGCMCFGLVN 371
Qy 230 -----ACLRCPVCVKIEVN-----TGSATARNVTVDN-----257
Db 372 GRCAQELPGRMTKMQCCCEPRGCMGIGITPEACPVRGSEYRRLCMLDGLPMGGIPGSAGS 431
Qy 258 -----PVPDGYSHASQORVLSFNLGDMRPGDKKVF---287
Db 432 RPGGTGGNGFAPSGNGYGGTGFPIPGNGGFSFGVGAGVGAGGQGIITGLTILN 491
Qy 288 -TVEFCPQRGQTNVATVYCGGHKGSAN-----VTVVNEPCVQVNISGA 333
Db 492 QTIDICKHANLCLNGRCIPTVSYRCECNMGYKQDANGDCIDVDECTSNPCTNGDCVNT 551
Qy 334 DWSYVCK-----PVEYSISVNSGDLVLHDVQDTLPSGVTVLEA 374
Db 552 PGSYKCHAGFORTPTKQACIDIDEICQNGVLCKN--GRCVNSDGSFQICNAGFEL--T 608
Qy 375 PGGEIC-----CNKVVMRIKEMCPGETLQPKLVVK-----AQVPGRFT 412
Db 609 TDGKNVDHDECITTNCLNGMCINEDGSFKICKPGFVLAPNGRYCTDVDECTPGICM 668
Qy 413 NQVAVTSESN--CGTCTSCAETHHWKGLAA-----THM---CVLDTNDPICV---455
Db 669 NGHCINSEGSFRC---DCP-----PGLAVGMDGRVCDVTHMRSTCYGGIKKGVCVRPF 718
Qy 456 -GENTVYRICVTN--RGSABEDTNVSLILKFSKELOPTASSG 493
Db 719 PGAVTKSECCANPDYGFGEPCQ-PCPAKNSAEFHGLCSSG 758

RESULT 13
DP87 D1CDI
ID DP87 D1CDI STANDARD; PRT; 555 AA.
AC Q04503:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Prespore protein Dp87 precursor.
GN DP87.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_Taxid:44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=94008533; PubMed=8404532;
RA Ozaki T., Nakao H., Orii H., Morio T., Takeuchi I., Tasaka M.;
RT "Developmental regulation of transcription of a novel
RL prespore-specific gene (Dp87) in Dictyostellium discoideum.";
RL Development 117:1299-1308(1993).
CC -!- SUBCELLULAR LOCATION: STORED IN PRESPORE VACUOLES UNTIL IT IS
CC DISCHARGED INTO THE INTERSPACE OF SPORES DURING SPORE FORMATION.
CC -!- INDUCTION: BY EXOGENOUS CAMP, REPRESSED BY DIF.
CC -!- SIMILARITY: CONTAINS 7 PRESPORE MOTIFS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC HBBP; D13973; BAA03083.1; -.
DR EMBL; P01542; LCCN.
DR DictyDb; DD03013;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR003645; FOLN.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00274; FOLN; 6.
KW Glycoprotein; Sporulation; Signal; Repeat.
```

```
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 555 PRESPORE PROTEIN DP87.
FT REPEAT 149 161 PRESPORE MOTIF 1.
FT REPEAT 176 188 PRESPORE MOTIF 2.
FT REPEAT 210 222 PRESPORE MOTIF 3.
FT REPEAT 241 253 PRESPORE MOTIF 4.
FT REPEAT 305 317 PRESPORE MOTIF 5.
FT REPEAT 333 345 PRESPORE MOTIF 6.
FT REPEAT 363 375 PRESPORE MOTIF 7.
FT DOMAIN 438 473 THR-RICH.
FT DOMAIN 476 555 ALA/SER-RICH.
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 555 AA; 58673 MW; 6D8A6DLC36F8DDD CRC64;

Query Match 3.8%; Score 113.5; DB 1; Length 555;
Best Local Similarity 17.4%; Pred. No. 0.36;
Matches 97; Conservative 66; Mismatches 215; Indels 179; Gaps 27;

Qy 75 CDKEFYPC-----EGRCPQVEAQOESCYSGLYSVKVNDCCNVEICSVPEYATVGSYPPI 130
Db 54 CKKEFF-CVNRDRHNCDDLSCAKNTRTGEIFEIWS-----CKPHRDFVYHSP---102
Qy 131 ETLAIGKDKCVDVVITQQLPCEA-----EFVSSDPETTTPTSDGKLVMKIDRLGAGDKCK1 185
Db 103 -----NTTTCESLGCCEARGMECEWVESPCYGTSCCPRIPCVGHGHHGKCDR 151
Qy 186 -----TVWKPLKEG-----CFTATVCA-----205
Db 152 MRCEGFYCEEQGSACVPHHDGCGNIQCPWGHYCVNEHGKRLVPHRPPRPVDPQCR 211
Qy 206 ---CPELRSYTKCGQPAICIKQEGPD---CACLRCPVCYKIEVVNTGSAIARNVTVDNPV 259
Db 212 NOHCPHGYSRVIKGCATCVRDARPHNLRCGFCGCEGSHCEVLEKHPVCVRNHRVPPHP 271
Qy 260 PDGYSHASQORVLSFNLGDMRPGDKKVFTEFCPQRGQITNVATVTCG-GHKCSANVT 318
Db 272 PP-----PQICGSVN-----CQPGYIC-----288
Qy 319 TVVN--EPCVQ-----VNISGADMSYVCKPVEYSISVNSPGDLVLHDVVIQDTLPSPVTV 371
Db 289 TIIINGHPTCIKRGDYLGNQTRCPHDYQCETISTINIVKCS-----KNDECKWHKCPPSSC 344
Qy 372 LEAPGGEIC-CNKVVMRIKEM--CPGETLQFKLV-----VKAQVPGRFTNQVAVTSES 422
Db 345 PNRNGPHCLANNVFPQICKVTCQPTD-FSCKMIRGNPTCIKARPP-----VPPPH 394
Qy 423 CGTCTSCAETHHWKGLAATHMCV-----LDTNDPICGENTVYRICV---TNRGSAED 473
Db 395 CSTCAELSSACNH-----VGMICIQVPSNCTNTRPCCPS---HPICIHPSITTAASTIA 445
Qy 474 TNVSLILKFSKELOPIASSGPKGTISGTVTVFDALPKLGSKESVEFSVTLKGIAPGDAR 533
Db 446 TTASTV-----ATTTSATTAGTTGTTGGTSDSSAASSADSSAA--SSSPSSA 495
Qy 534 GEAILSSDTLTSPVSDT 550
Db 496 ASSAASSEPPSSSAASS 512

RESULT 14
ID TENA_CHICK STANDARD; PRT; 1808 AA.
AC P10039; P13132; 073584; 073585;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)
DE (GMEM) (J1) (Mitotendinous antigen) (Glioma-associated-extracellular
DE matrix antigen) (GP 150-225).
OS Gallus gallus (Chicken).
```





|    |          |          |            |                                    |
|----|----------|----------|------------|------------------------------------|
| FT | DISULFID | 533      | 543        | BY SIMILARITY.                     |
| FT | DISULFID | 537      | 548        | BY SIMILARITY.                     |
| FT | DISULFID | 550      | 559        | BY SIMILARITY.                     |
| FT | DISULFID | 564      | 574        | BY SIMILARITY.                     |
| FT | DISULFID | 568      | 579        | BY SIMILARITY.                     |
| FT | DISULFID | 581      | 590        | BY SIMILARITY.                     |
| FT | CARBOHYD | 38       | 38         | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 168      | 168        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 186      | 186        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 328      | 328        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 603      | 603        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 643      | 643        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 751      | 751        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 759      | 759        | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1050     | 1050       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1090     | 1090       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1101     | 1101       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1112     | 1112       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1153     | 1153       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1183     | 1183       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1416     | 1416       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1736     | 1736       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | CARBOHYD | 1769     | 1769       | N-LINKED (GLCNAC. .) (POTENTIAL).  |
| FT | VARSPLIC | 1043     | 1224       | MISSING (IN 200 KDA ISOFORM).      |
| FT | VARSPLIC | 1043     | 1315       | MISSING (IN 190 KDA ISOFORM).      |
| FT | CONFLICT | 182      | 181        | W -> R (IN REF. 2).                |
| FT | CONFLICT | 563      | 571        | SPCNDNNV -> PAPMTATTW (IN REF. 3). |
| FT | CONFLICT | 598      | 598        | E -> G (IN REF. 3).                |
| FT | CONFLICT | 838      | 838        | T -> TEY (IN REF. 3).              |
| FT | CONFLICT | 886      | 886        | N -> F (IN REF. 3).                |
| FT | SEQUENCE | 1808 AA; | 198858 MW; | B924A06CF9ED6DE CRC64;             |

Query Match 3.8%; Score 113.5; DB 1; Length 1808;

Best Local Similarity 21.0%; Pred. No. 1.4;

Matches 127; Conservative 64; Mismatches 200; Indels 213; Gaps

|    |     |  |  |                               |
|----|-----|--|--|-------------------------------|
| QY | 82  | CEEGRCQPVAAQESCYGRLYSVKVNDDCNVEIC                                  | -----VITQQLPCEAEFVSSOPETPTTSDGKLWVKIDRL 177            | -----QSVPEYATVGSYP 128        |
| DB | 357 | CENGLC   | -----VCHEGF-----VGDDCSQKPCPKDCNNRGHCYDGRVCHEGYLG       | -----401                      |
| QY | 129 | PIEILAIGKKDCVDV  | -----VITQQLPCEAEFVSSOPETPTTSDGKLWVKIDRL 177            | -----CGEL 438                 |
| DB | 402 | -----EDCGELRCPDCHNHRGRCINGQCVCDESGFED                              | -----CGEL 438  |                               |
| QY | 178 | GAGDKCKITVWVKPLKEGCGFFAATVC  | -----ACPELRSTYKCGQPAICIKQE                             | -----225                      |
| DB | 439 | RCPDNDCH   | -----NRGRVCVNGQCCEHGFEGEDCGELRCPDCHNHRGRCVNGQCVCDE 490 |                               |
| QY | 226 | ---GPDCACLRCP-VCYKIEVNTGSAIARNVTVDNPPDGY                           | -----SHASG 268   |                               |
| DB | 491 | GYTGEDCGELRCPDNDCH   | -----NRGRVCVNGQCVCDE                                   | -----GFMGEDCGELSCPNDCHQHG 541 |
| QY | 269 | QRVLSFNLGDMRPGDKKKVFTFEPCPQRRGQITNVAITVYCGGHKGSANVTYVNVNEPCV-Q 327 |  |                               |
| DB | 542 | RCV-----DGRVCVCEHGFEGEDCGRS  | -----C-PDNCN-NVGRVCVCEGRVCE 582                        |                               |
| QY | 328 | VNISGADWSVCKPVEYSIVSNPGDLVL  | -----HDVVI-----QDTL 365                                |                               |
| DB | 583 | EGYMGIDCSDVSPPTB--LTVTNVTDKTVNLEWKHENLVNLEYLVTVPTSSGGDLQFTV 640    |  |                               |
| QY | 366 | P-----SGVTVLE-APGGE-----ICCNK-----VWVRKEMCPG-ETLOFKLVVKAQVPG 409   |  |                               |
| DB | 641 | PGNQTSAITHELEPGVEYFIRVFAILKNKSTPVSARVATYUPAPEGLKFKSVRETSVQV 700    |  |                               |
| QY | 410 | R-----FTNOVAVTSESNCGTCTSCAETTTTHWKGLA-----ATHWCVID 448             |  |                               |
| DB | 701 | EMDPLSISFDGWEIVFRNMOKKDDNGDITSSLKRPETSYMQGPLAPGOQYNVSLHIVRNN 760   |  |                               |
| QY | 449 | TNDPICVGENITY--RICVTNRSRSED-TNWSLLIKFSELOPIASSGPTKG--TISGNT 503    |  |                               |
| DB | 761 | TRGP---GLSRVITTKLDAPSQIEAKVDITDALTITWKSPLAEIEGIEIUTYGPKDVPGR 817   |  |                               |

|    |          |      |      |                                  |    |          |      |      |                |
|----|----------|------|------|----------------------------------|----|----------|------|------|----------------|
| FT | SIGNAL   | 1    | 28   | POTENTIAL.                       | FT | DISULFID | 514  | 526  | BY SIMILARITY. |
| FT | CHAIN    | 29   | 2907 | FIBRILLIN 2.                     | FT | DISULFID | 532  | 542  | BY SIMILARITY. |
| FT | DOMAIN   | 111  | 142  | EGF-LIKE 1, NON-CALCIUM BINDING. | FT | DISULFID | 537  | 551  | BY SIMILARITY. |
| FT | DOMAIN   | 145  | 176  | EGF-LIKE 2, NON-CALCIUM BINDING. | FT | DISULFID | 553  | 566  | BY SIMILARITY. |
| FT | DOMAIN   | 176  | 208  | EGF-LIKE 3, NON-CALCIUM BINDING. | FT | DISULFID | 572  | 584  | BY SIMILARITY. |
| FT | DOMAIN   | 276  | 317  | EGF-LIKE 4, CALCIUM-BINDING.     | FT | DISULFID | 579  | 593  | BY SIMILARITY. |
| FT | DOMAIN   | 318  | 359  | EGF-LIKE 5, CALCIUM-BINDING.     | FT | DISULFID | 595  | 608  | BY SIMILARITY. |
| FT | REPEAT   | 360  | 426  | TGFBP 1.                         | FT | DISULFID | 614  | 625  | BY SIMILARITY. |
| FT | DOMAIN   | 487  | 527  | EGF-LIKE 6, NON-CALCIUM BINDING. | FT | DISULFID | 620  | 634  | BY SIMILARITY. |
| FT | DOMAIN   | 528  | 567  | EGF-LIKE 7, CALCIUM-BINDING.     | FT | DISULFID | 636  | 649  | BY SIMILARITY. |
| FT | DOMAIN   | 568  | 609  | EGF-LIKE 8, CALCIUM-BINDING.     | FT | DISULFID | 655  | 666  | BY SIMILARITY. |
| FT | DOMAIN   | 610  | 650  | EGF-LIKE 9, CALCIUM-BINDING.     | FT | DISULFID | 661  | 675  | BY SIMILARITY. |
| FT | DOMAIN   | 651  | 691  | EGF-LIKE 10, CALCIUM-BINDING.    | FT | DISULFID | 677  | 690  | BY SIMILARITY. |
| FT | REPEAT   | 692  | 760  | TGFBP 2.                         | FT | DISULFID | 765  | 777  | BY SIMILARITY. |
| FT | DOMAIN   | 761  | 802  | EGF-LIKE 11, CALCIUM-BINDING.    | FT | DISULFID | 772  | 786  | BY SIMILARITY. |
| FT | DOMAIN   | 803  | 844  | EGF-LIKE 12, CALCIUM-BINDING.    | FT | DISULFID | 788  | 801  | BY SIMILARITY. |
| FT | DOMAIN   | 845  | 883  | EGF-LIKE 13, CALCIUM-BINDING.    | FT | DISULFID | 807  | 819  | BY SIMILARITY. |
| FT | DOMAIN   | 948  | 989  | EGF-LIKE 14, CALCIUM-BINDING.    | FT | DISULFID | 814  | 828  | BY SIMILARITY. |
| FT | REPEAT   | 990  | 1065 | TGFBP 3.                         | FT | DISULFID | 830  | 843  | BY SIMILARITY. |
| FT | DOMAIN   | 1066 | 1107 | EGF-LIKE 15, CALCIUM-BINDING.    | FT | DISULFID | 849  | 859  | BY SIMILARITY. |
| FT | DOMAIN   | 1108 | 1150 | EGF-LIKE 16, CALCIUM-BINDING.    | FT | DISULFID | 854  | 868  | BY SIMILARITY. |
| FT | DOMAIN   | 1151 | 1192 | EGF-LIKE 17, CALCIUM-BINDING.    | FT | DISULFID | 870  | 883  | BY SIMILARITY. |
| FT | DOMAIN   | 1193 | 1234 | EGF-LIKE 18, CALCIUM-BINDING.    | FT | DISULFID | 952  | 964  | BY SIMILARITY. |
| FT | DOMAIN   | 1235 | 1275 | EGF-LIKE 19, CALCIUM-BINDING.    | FT | DISULFID | 959  | 973  | BY SIMILARITY. |
| FT | DOMAIN   | 1276 | 1317 | EGF-LIKE 20, CALCIUM-BINDING.    | FT | DISULFID | 975  | 988  | BY SIMILARITY. |
| FT | DOMAIN   | 1318 | 1359 | EGF-LIKE 21, CALCIUM-BINDING.    | FT | DISULFID | 1070 | 1082 | BY SIMILARITY. |
| FT | DOMAIN   | 1360 | 1400 | EGF-LIKE 22, CALCIUM-BINDING.    | FT | DISULFID | 1077 | 1091 | BY SIMILARITY. |
| FT | DOMAIN   | 1401 | 1441 | EGF-LIKE 23, CALCIUM-BINDING.    | FT | DISULFID | 1093 | 1106 | BY SIMILARITY. |
| FT | DOMAIN   | 1442 | 1483 | EGF-LIKE 24, CALCIUM-BINDING.    | FT | DISULFID | 1112 | 1124 | BY SIMILARITY. |
| FT | DOMAIN   | 1484 | 1524 | EGF-LIKE 25, CALCIUM-BINDING.    | FT | DISULFID | 1119 | 1133 | BY SIMILARITY. |
| FT | DOMAIN   | 1525 | 1565 | EGF-LIKE 26, CALCIUM-BINDING.    | FT | DISULFID | 1135 | 1149 | BY SIMILARITY. |
| FT | REPEAT   | 1566 | 1642 | TGFBP 4.                         | FT | DISULFID | 1155 | 1167 | BY SIMILARITY. |
| FT | DOMAIN   | 1643 | 1684 | EGF-LIKE 27, CALCIUM-BINDING.    | FT | DISULFID | 1162 | 1176 | BY SIMILARITY. |
| FT | DOMAIN   | 1685 | 1726 | EGF-LIKE 28, CALCIUM-BINDING.    | FT | DISULFID | 1178 | 1191 | BY SIMILARITY. |
| FT | REPEAT   | 1727 | 1800 | TGFBP 5.                         | FT | DISULFID | 1197 | 1209 | BY SIMILARITY. |
| FT | DOMAIN   | 1801 | 1842 | EGF-LIKE 29, CALCIUM-BINDING.    | FT | DISULFID | 1204 | 1218 | BY SIMILARITY. |
| FT | DOMAIN   | 1843 | 1884 | EGF-LIKE 30, CALCIUM-BINDING.    | FT | DISULFID | 1220 | 1233 | BY SIMILARITY. |
| FT | DOMAIN   | 1885 | 1926 | EGF-LIKE 31, CALCIUM-BINDING.    | FT | DISULFID | 1239 | 1250 | BY SIMILARITY. |
| FT | DOMAIN   | 1927 | 1965 | EGF-LIKE 32, CALCIUM-BINDING.    | FT | DISULFID | 1246 | 1259 | BY SIMILARITY. |
| FT | DOMAIN   | 1966 | 2008 | EGF-LIKE 33, CALCIUM-BINDING.    | FT | DISULFID | 1261 | 1274 | BY SIMILARITY. |
| FT | DOMAIN   | 2009 | 2048 | EGF-LIKE 34, CALCIUM-BINDING.    | FT | DISULFID | 1280 | 1292 | BY SIMILARITY. |
| FT | DOMAIN   | 2049 | 2090 | EGF-LIKE 35, CALCIUM-BINDING.    | FT | DISULFID | 1287 | 1301 | BY SIMILARITY. |
| FT | REPEAT   | 2091 | 2163 | TGFBP 6.                         | FT | DISULFID | 1303 | 1316 | BY SIMILARITY. |
| FT | DOMAIN   | 2164 | 2205 | EGF-LIKE 36, CALCIUM-BINDING.    | FT | DISULFID | 1322 | 1334 | BY SIMILARITY. |
| FT | DOMAIN   | 2206 | 2245 | EGF-LIKE 37, CALCIUM-BINDING.    | FT | DISULFID | 1329 | 1343 | BY SIMILARITY. |
| FT | DOMAIN   | 2246 | 2286 | EGF-LIKE 38, CALCIUM-BINDING.    | FT | DISULFID | 1345 | 1358 | BY SIMILARITY. |
| FT | DOMAIN   | 2287 | 2330 | EGF-LIKE 39, CALCIUM-BINDING.    | FT | DISULFID | 1364 | 1377 | BY SIMILARITY. |
| FT | DOMAIN   | 2331 | 2372 | EGF-LIKE 40, CALCIUM-BINDING.    | FT | DISULFID | 1371 | 1386 | BY SIMILARITY. |
| FT | REPEAT   | 2373 | 2441 | TGFBP 7.                         | FT | DISULFID | 1388 | 1399 | BY SIMILARITY. |
| FT | DOMAIN   | 2442 | 2483 | EGF-LIKE 41, CALCIUM-BINDING.    | FT | DISULFID | 1405 | 1418 | BY SIMILARITY. |
| FT | DOMAIN   | 2484 | 2524 | EGF-LIKE 42, CALCIUM-BINDING.    | FT | DISULFID | 1412 | 1427 | BY SIMILARITY. |
| FT | DOMAIN   | 2525 | 2563 | EGF-LIKE 43, CALCIUM-BINDING.    | FT | DISULFID | 1429 | 1440 | BY SIMILARITY. |
| FT | DOMAIN   | 2564 | 2606 | EGF-LIKE 44, CALCIUM-BINDING.    | FT | DISULFID | 1446 | 1458 | BY SIMILARITY. |
| FT | DOMAIN   | 2607 | 2646 | EGF-LIKE 45, CALCIUM-BINDING.    | FT | DISULFID | 1453 | 1467 | BY SIMILARITY. |
| FT | DOMAIN   | 2647 | 2687 | EGF-LIKE 46, CALCIUM-BINDING.    | FT | DISULFID | 1469 | 1482 | BY SIMILARITY. |
| FT | DOMAIN   | 2688 | 2727 | EGF-LIKE 47, CALCIUM-BINDING.    | FT | DISULFID | 1488 | 1499 | BY SIMILARITY. |
| FT | DISULFID | 115  | 124  | BY SIMILARITY.                   | FT | DISULFID | 1494 | 1508 | BY SIMILARITY. |
| FT | DISULFID | 119  | 130  | BY SIMILARITY.                   | FT | DISULFID | 1510 | 1523 | BY SIMILARITY. |
| FT | DISULFID | 132  | 141  | BY SIMILARITY.                   | FT | DISULFID | 1529 | 1540 | BY SIMILARITY. |
| FT | DISULFID | 149  | 159  | BY SIMILARITY.                   | FT | DISULFID | 1549 | 1564 | BY SIMILARITY. |
| FT | DISULFID | 153  | 164  | BY SIMILARITY.                   | FT | DISULFID | 1551 | 1564 | BY SIMILARITY. |
| FT | DISULFID | 166  | 175  | BY SIMILARITY.                   | FT | DISULFID | 1647 | 1659 | BY SIMILARITY. |
| FT | DISULFID | 180  | 190  | BY SIMILARITY.                   | FT | DISULFID | 1654 | 1668 | BY SIMILARITY. |
| FT | DISULFID | 184  | 196  | BY SIMILARITY.                   | FT | DISULFID | 1670 | 1683 | BY SIMILARITY. |
| FT | DISULFID | 198  | 207  | BY SIMILARITY.                   | FT | DISULFID | 1689 | 1701 | BY SIMILARITY. |
| FT | DISULFID | 280  | 292  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 307  | 301  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 303  | 316  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 322  | 334  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 329  | 343  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 345  | 358  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 491  | 503  | BY SIMILARITY.                   |    |          |      |      |                |
| FT | DISULFID | 498  | 512  | BY SIMILARITY.                   |    |          |      |      |                |

Query Match

Best Local Similarity 3.8%; Score 113; D

Matches 107; Conservative 54; Mismatches

62

NKQPVQKRGARFCDEKFEYPCEEGR-----CQPVQEAQ

221

NNQMCQQLTGIVCTKTLCATIGRAWGHPCEMCPAQ

QY

DB



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

DM protein - protein search, using sw model

Run on: May 25, 2002, 22:15:37 ; Search time 95.87 Seconds  
(without alignments)

1003.288 Million cell updates/sec

Title: US-09-523-647-2

Perfect score: 2951

Sequence: 1 MSKLIRRVTVTLTSMASCSFASGGTEAAVAESLIITKIVASAEKTPAPV-PMTAKKVRV 556

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID     | Description        |
|------------|--------|-------------|--------|--------|--------------------|
| 1          | 2523.5 | 85.5        | 557    | Q9AIS7 | Q9ais7 chlamydophi |
| 2          | 2488   | 84.3        | 558    | P94664 | P94664 chlamydophi |
| 3          | 2418.5 | 82.0        | 534    | Q9ZAW1 | Q9zaw1 chlamydophi |
| 4          | 2323.5 | 78.7        | 534    | Q9ZAW0 | Q9zaw0 chlamydophi |
| 5          | 2176   | 73.7        | 554    | Q9PJY0 | Q9pjy0 chlamydia m |
| 6          | 2157.5 | 73.1        | 547    | Q933I7 | Q933i7 chlamydia t |
| 7          | 2127   | 72.1        | 395    | Q956B3 | Q956b3 chlamydia p |
| 8          | 2008.5 | 68.1        | 511    | Q93FV8 | Q93fv8 chlamydia t |
| 9          | 1917   | 65.0        | 395    | Q9X4I2 | Q9x4i2 chlamydophi |
| 10         | 1833   | 62.1        | 395    | Q9X4I1 | Q9x4i1 chlamydophi |
| 11         | 956    | 32.4        | 183    | Q46164 | Q46164 chlamydia p |
| 12         | 920    | 31.2        | 176    | Q9AG10 | Q9ag10 chlamydia p |
| 13         | 685    | 23.2        | 182    | Q9AF83 | Q9af83 chlamydophi |
| 14         | 674    | 22.8        | 128    | Q9RB53 | Q9rb53 chlamydia p |
| 15         | 669    | 22.7        | 178    | Q46163 | Q46163 chlamydophi |
| 16         | 659    | 22.3        | 178    | Q46165 | Q46165 chlamydophi |

|    |       |      |      |    |        |                    |
|----|-------|------|------|----|--------|--------------------|
| 17 | 578   | 19.6 | 157  | 2  | Q9AF85 | Q9af85 chlamydophi |
| 18 | 578   | 19.6 | 157  | 2  | Q9AF84 | Q9af84 chlamydophi |
| 19 | 509   | 17.2 | 173  | 2  | Q9AF82 | Q9af82 chlamydia s |
| 20 | 151.5 | 5.1  | 1474 | 17 | O27146 | O27146 methanother |
| 21 | 135   | 4.6  | 2233 | 5  | Q94711 | Q94711 paramecium  |
| 22 | 130.5 | 4.4  | 756  | 16 | Q9RZS5 | Q9rzs5 deinococcus |
| 23 | 129.5 | 4.4  | 667  | 5  | Q95WU1 | Q95wu1 giardia lam |
| 24 | 129.5 | 4.4  | 757  | 5  | Q9VZF2 | Q9vzf2 drosophila  |
| 25 | 129   | 4.4  | 556  | 5  | Q9NGZ3 | Q9ngz3 giardia lam |
| 26 | 128.5 | 4.4  | 1404 | 5  | Q9VB65 | Q9vb65 drosophila  |
| 27 | 127.5 | 4.3  | 998  | 16 | Q9RWI2 | Q9rwi2 deinococcus |
| 28 | 127.5 | 4.3  | 2327 | 13 | Q9IBG7 | Q9ibg7 xenopus lae |
| 29 | 127   | 4.3  | 693  | 5  | O01600 | O01600 caenorhabdi |
| 30 | 125.5 | 4.3  | 3198 | 5  | Q9U8G8 | Q9u8g8 manduca sex |
| 31 | 124.5 | 4.2  | 739  | 5  | Q9GS24 | Q9gs24 giardia lam |
| 32 | 124   | 4.2  | 2440 | 5  | Q9GP97 | Q9gp97 caenorhabdi |
| 33 | 123.5 | 4.2  | 765  | 5  | Q9NL50 | Q9nl50 sarcophaga  |
| 34 | 123.5 | 4.2  | 2174 | 5  | Q9GQR0 | Q9gqr0 drosophila  |
| 35 | 123.5 | 4.2  | 2906 | 11 | Q9WU99 | Q9wuh9 rattus norv |
| 36 | 122.5 | 4.2  | 950  | 13 | Q90Z44 | Q90z44 gallus gall |
| 37 | 122   | 4.1  | 491  | 16 | Q9RY09 | Q9ry09 deinococcus |
| 38 | 122   | 4.1  | 955  | 4  | Q96DN2 | Q96dn2 homo sapien |
| 39 | 122   | 4.1  | 2910 | 11 | O55225 | O55225 mus musculu |
| 40 | 122   | 4.1  | 2931 | 5  | Q9W2C6 | Q9w2c6 drosophila  |
| 41 | 121.5 | 4.1  | 2809 | 4  | Q96JP8 | Q96jp8 homo sapien |
| 42 | 121   | 4.1  | 641  | 17 | O58295 | O58295 pyrococcus  |
| 43 | 121   | 4.1  | 1546 | 4  | Q9NS27 | Q9ns27 homo sapien |
| 44 | 121   | 4.1  | 1551 | 4  | O75445 | O75445 homo sapien |
| 45 | 120.5 | 4.1  | 354  | 5  | Q9XV21 | Q9xv21 caenorhabdi |

ALIGNMENTS

RESULT 1

|        |   |              |      |         |
|--------|---|--------------|------|---------|
| Q9AIS7 | Q9AIS7  | PRELIMINARY; | PRT; | 557 AA. |
| ID     | Q9AIS7  |              |      |         |
| AC     | Q9AIS7;   |              |      |         |
| DT     | 01-JUN-2001 (TREMBlrel. 17, Created)                  |              |      |         |
| DT     | 01-JUN-2001 (TREMBlrel. 17, Last sequence update)     |              |      |         |
| DT     | 01-OCT-2001 (TREMBlrel. 18, Last annotation update)   |              |      |         |
| DE     | 60 KDA CYSTEINE-RICH MEMBRANE COMPLEX PROTEIN.        |              |      |         |
| GN     | CMCB.   |              |      |         |
| OS     | Chlamydophila abortus.                                |              |      |         |
| OC     | Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila. |              |      |         |
| OX     | NCBI_TaxID=893555;                                    |              |      |         |
| RN     | [1]   |              |      |         |
| RP     | SEQUENCE FROM N.A.                                    |              |      |         |
| RC     | STRAIN=EBA;   |              |      |         |
| RX     | MEDLINE=21078680; PubMed=11211261;                    |              |      |         |
| RA     | Bush R.M., Everett K.D.;                              |              |      |         |
| RT     | "Molecular evolution of the Chlamydiaceae."           |              |      |         |
| RL     | Int. J. Syst. Evol. Microbiol. 51:203-220(2001).      |              |      |         |
| DR     | EMBL; AF240773; AAG60550.1; -                         |              |      |         |
| DR     | InterPro: IPR003506; Chlam_OMP6.                      |              |      |         |
| DR     | PRINTS: PR01336; CHLAMIDIAOM6.                        |              |      |         |
| SQ     | SEQUENCE 557 AA; 59761 MW; 817A20B071AD9609 CRC64;    |              |      |         |

|                       |                  |   |           |             |
|-----------------------|------------------|---|-----------|-------------|
| Query Match           | 85.5%            | Score 2523.5;   | DB 2;     | Length 557; |
| Best Local Similarity | 84.4%            | Pred. No. 3.6e-198;   |           |             |
| Matches 470;          | Conservative 33; | Mismatches 53;  | Indels 1; | Gaps 1;     |
| QY                    | 1                | MSKLIRRVTVTLTSMASCSFASGGTEAAVAESLIITKIVASAEKTPAPV-PMTAKKVRV | 59        |             |
| Db                    | 1                | MSKLIRRVTVTLTSMASCSFASGGTEAAVAESLIITKIVASAEKTPAPV-PMTAKKVRV | 60        |             |
| QY                    | 60               | RRNKQPEQKSRGAFCDKEFPCEGRCPQVEAQGCYGRLYSVKVNDDCNVEICQSV      | 119       |             |
| Db                    | 61               | RNRKQPEQKSRGAFCDKEFPCEGRCPQVEAQGCYGRLYSVKVNDDCNVEICQSV      | 120       |             |
| QY                    | 120              | EYATVGSPPYETIATIGKKDCVDVITFQQLPCBAEFVSSDPETPTSDGKLVKIDRLGA  | 179       |             |

Db 121 EYATVGSPPYPIEILAVGKKDCVNVVITQOLPCEVEFVSDDPATPTSDSKLIWTIDCLGQ 180  
QY 180 GDKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 239  
Db 181 QGERCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 240  
Db 181 GEKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 240  
QY 240 IEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVTFEFCPQRRGOI 299  
Db 241 IEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVTFEFCPQRRGOI 299  
QY 300 TNVATVYCGGHKCSANVTTVVNEPCVQVNSGADWSYCKPVEYSISVSNPGDLVLHDV 359  
Db 301 TNVATVYCGGHKCSANVTTVVNEPCVQVNSGADWSYCKPVEYSISVSNPGDLVLHDV 359  
Db 301 TNVATVYCGGHKCSANVTTVVNEPCVQVNSGADWSYCKPVEYSISVSNPGDLVLHDV 360  
QY 360 VIQDTPSGVTVLEAPGGEICCNKVMWRIKEMCPGETLQFKLVKVAQVPGRETNOVAVTS 419  
Db 361 VVEDTPSGATILEAPGAEICCNKAVWCIKEMCPGETLQFKLVKVAQVPGRETNOVAVTS 419  
QY 420 ESNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLI 479  
Db 421 NSDCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLI 480  
QY 480 LKFSKELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVESVTLKGTPAGDARGEAIL 539  
Db 481 LKFSKELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVESVTLKGTPAGDARGEAIL 540  
QY 540 SDTLTSPVSDTENTHVV 556  
Db 541 SDTLTVPVADTENTHVV 557

RESULT 2  
P94664  
ID P94664 PRELIMINARY; PRT; 558 AA.  
AC P94664;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE OUTER MEMBRANE PROTEIN 2.  
GN OMP2  
OS Chlamydomophila caviae.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
OX NCBI\_TaxID=83557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS (GPIC);  
RX MEDLINE=97075924; PubMed=8918247;  
RA Hsia R.C., Bavoil P.M.;  
RT "Sequence analysis of the omp2 region of Chlamydia psittaci strain  
GPIC: structural and functional implications.";  
RL Gene 176:155-162(1996).  
DR EMBL; U41759; AAB41143.1; -;  
DR InterPro: IPR003506; ChlamOMP6.  
DR PRINTS: PR01336; CHLAMIDIAOM6.  
SQ SEQUENCE 558 AA; 60213 MW; 452E2BD79078935C CRC64;

Query Match 84.3%; Score 2488; DB 2; Length 558;  
Best Local Similarity 82.1%; Pred. No. 2.9e-195;  
Matches 458; Conservative 44; Mismatches 54; Indels 2; Gaps 2;  
QY 1 MSLKRLRVVTVLALSMASCFASGGIEAAVAESLITKIVASAEETKPAPVPM-TAKKRVLV 59  
Db 1 MSLKRLRVVTVLALSMASCFASGGIEAAVAESLITKIVASAEETKPAPVPM-TAKKRVLV 59  
QY 60 RNNKQPVQKSGARCDKEFEYPCBEGRCQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSV 118  
Db 61 RNNKQPVQKSGARCDKEFEYPCBEGRCQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSV 118  
QY 119 PEYATVGSPPYPIEILAVGKKDCVNVVITQOLPCEVEFVSDDPATPTSDSKLIWTIDCLG 178  
Db 121 PEYATVGSPPYPIEILAVGKKDCVNVVITQOLPCEVEFVSDDPATPTSDSKLIWTIDCLG 180

QY 179 AGDKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCY 238  
Db 181 QGERCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCY 240  
QY 239 KIEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVTFEFCPQRRGO 298  
Db 241 KIEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVTFEFCPQRRGO 298  
QY 299 ITNVATVYCGGHKCSANVTTVVNEPCVQVNSGADWSYCKPVEYSISVSNPGDLVLHD 358  
Db 301 ITNVATVYCGGHKCSANVTTVVNEPCVQVNSGADWSYCKPVEYSISVSNPGDLVLHD 360  
QY 359 VIQDTPSGVTVLEAPGGEICCNKVMWRIKEMCPGETLQFKLVKVAQVPGRETNOVAVT 418  
Db 361 VVEDTPSGATILEAPGAEICCNKAVWCIKEMCPGETLQFKLVKVAQVPGRETNOVAVT 420  
QY 419 SESNCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTVNRGSAEDTNVSL 478  
Db 421 TNSDCGTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTVNRGSAEDTNVSL 480  
QY 479 ILKFSKELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVESVTLKGTPAGDARGEAIL 538  
Db 481 ILKFSKELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVESVTLKGTPAGDARGEAIL 540  
QY 539 SSDTLTSPVSDTENTHVV 556  
Db 541 SSDTLTVPVADTENTHVV 558

RESULT 3  
Q9ZAW1  
ID Q9ZAW1 PRELIMINARY; PRT; 534 AA.  
AC Q9ZAW1;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).  
OS Chlamydomophila abortus.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
OX NCBI\_TaxID=83555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S26/3;  
RA Sheehy N., Markey B., Quinn P.J.;  
RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes.";  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U76760; AAD09597.1; -;  
DR InterPro: IPR003506; ChlamOMP6.  
DR PRINTS: PR01336; CHLAMIDIAOM6.  
FT NON\_TER 1  
FT NON\_TER 534  
SQ SEQUENCE 534 AA; 57137 MW; 2789820B041E92BE CRC64;

Query Match 82.0%; Score 2418.5; DB 2; Length 534;  
Best Local Similarity 84.1%; Pred. NO. 1.3e-189;  
Matches 449; Conservative 32; Mismatches 52; Indels 1; Gaps 1;  
QY 10 TVLALTSMASSCFASGGIEAAVAESLITKIVASAEETKPAPV-PMTAKKRVLRNKPQVBEQ 68  
Db 1 TVLALTSMASSCFASGGIEAAVAESLITKIVASAEETKPAPV-PMTAKKRVLRNKPQVBEQ 68  
QY 69 KSRGAFCDKEFEYPCBEGRCQVPEAQOESCYGRLYSVKVNDDCNVEICQSVPEYATVGS 128  
Db 61 KHTGAFCDKEFEYPCBEGRCQVPEAQOESCYGRLYSVKVNDDCNVEICQSVPEYATVGS 120  
QY 129 PIETLAIKGGKDCVNVVITQOLPCEVEFVSDDPATPTSDSKLIWTIDCLGADCKKITVW 188  
Db 121 PIETLAIKGGKDCVNVVITQOLPCEVEFVSDDPATPTSDSKLIWTIDCLGADCKKITVW 180  
QY 189 VKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKIEVNTGSA 248  
Db 181 VKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKIEVNTGSA 240

```
QY 249 IARNVTVNDPVDGYSHASGQVLSFNLDMPGDKKVFYVEFCQRRQITNATVYC 308
Db 241 IARNVVNDPVDGYTHASGQVLSFNLDMPGDKKVFYVEFCQRRQITNATVYC 300
QY 309 GGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHVDVYIOTLPSG 368
Db 301 GGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYTIIVSNLGLDKLYDVVVDVPSG 360
QY 369 VTVLEAPGECICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSSESNGTCTS 428
Db 361 ATILEAEGAEICCNKAVWCIREKCPGETLQFKLVVKAQSPGRTNOVAVTSSESNGTCTS 420
QY 429 CAETTTHWKGAAATHMVCVLDNDPICVGVNTVYRICVTVNRGSAEDTNVSLILKFSKELQP 488
Db 421 CAETTTHWKGAAATHMVCVLDNDPICVGVNTVYRICVTVNRGSAEDTNVSLILKFSKELQP 480
QY 489 IASSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 542
Db 481 VSSGPTKGTITGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 534

RESULT 4
ID Q9ZAW0 PRELIMINARY; PRT; 534 AA.
AC Q9ZAW0;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydia pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W73;
RA Sheehy N., Markey B., Quinn P.J.;
RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U76761; AAD09598.1;
DR InterPro: IPR003506; Chlam_OMP6.
DR PRINTS: PR01336; CHLAMIDIOM6.
FT NON_TER 1
FT NON_TER 534
SQ SEQUENCE 534 AA; 57298 MW; 39299DBD6A8534258 CRC64;
```

---

```
QY 10 TVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPTAKKVLVRRNKQPV-- 67
Db 1 TVLALTSMTSFASGGIEAAVAESLITKIVANAESKRA---ADHKKEAKHNKKHKSAT 57
QY 68 ---QKSRGAFCDKEFPCEEGRCQPVQAQOESCYGRLYSVKVNDDCNVETOSVPEYATV 124
Db 58 KHSRNSPSCDKEF-PCEKNQQQVESRQESCFCGRMYSIKYVNDNDCNVELSQSVPEYATV 116
QY 125 GSPYPIEIIAIGKCKDVVITQQLPCEAEFVSSDPTTSDGKLWKIDRLGAGDKCK 184
Db 117 GSPYPIEIIAIGKCKDVVITQQLPCEAEFVSSDPTTADGKLWKIDRLGQGERCK 176
QY 185 ITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICIKOEGPDCACLRCPVCYKIEVNV 244
Db 177 ITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICIKOEGPDCACLRCPVCYKIEVNV 236
QY 245 TGSIAIRNVTVNDPVDGYSHASGQVLSFNLDMPGDKKVFYVEFCQRRQITNATV 304
Db 237 TGSIAIRNVTVNDPVDGYSHASGQVLSFNLDMPGDKKVFYVEFCQRRQITNATV 296
QY 305 VTYCGGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHVDVYIOTL 364
Db 301 GGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHVDVYIOTL 358
```

```
Db 297 VSYCGGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYTIIVSNPGDLVLRNVVVD 356
QY 365 LPSGVTVLEAPGECICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRTNOVAVTSSESNG 424
Db 357 LPSGALILEAAGAEICCNKAVWCIREKCPGETLQFKLVVKAQVPGRTNOVAVTSSESNG 416
QY 425 TCTSAETTTTHWKGAAATHMVCVLDNDPICVGVNTVYRICVTVNRGSAEDTNVSLILKFSK 484
Db 417 VCTSCADATTYWKGLAATHMVCVLDNDPICVGVNTVYRICVTVNRGSAEDTNVSLILKFSK 476
QY 485 ELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 542
Db 477 ELQPISSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 534

RESULT 5
ID Q9PJUV0 PRELIMINARY; PRT; 554 AA.
AC Q9PJUV0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 60 KDA OUTER MEMBRANE PROTEIN.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL Neumoniae AR39.";
DR EMBL: AE002341; AAF39537.1;
DR TIGR: TC0727;
KW Complete proteome.
SQ SEQUENCE 554 AA; 59784 MW; 1CF81B2471C27FD4 CRC64;
```

---

```
Query Match 73.7%; Score 2176; DB 16; Length 554;
Best Local Similarity 72.8%; Pred. No. 9.7e-170;
Matches 405; Conservative 74; Mismatches 69; Indels 8; Gaps 3;

QY 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPTAKKVLVRRNKQ 60
Db 7 MNKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPTAKKVLVRRNKQ 65
QY 61 RNKQPVQKSRGAFCDKEFPCEEGRCQPVQAQOESCYGRLYSVKVNDDCNVETOSVPE 120
Db 66 KKHQ-----NRTSVVRKEVTAVRDTRK--VEPRQDSCFCGKMYTKVNDNDNRVETOSVPE 118
QY 121 YATVGSYPYPIEIIAIGKCKDVVITQQLPCEAEFVSSDPTTSDGKLWKIDRLGAG 180
Db 119 YATVGSYPYPIEIIAIGKCKDVVITQQLPCEAEFVSSDPTTADGKLWKIDRLGAG 178
QY 181 DKCKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICIKOEGPDCACLRCPVCYKI 240
Db 179 ESKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICIKOEGPDCACLRCPVCYKI 238
QY 241 EVVNTGSAIRNVTVNDPVDGYSHASGQVLSFNLDMPGDKKVFYVEFCQRRQIT 300
Db 239 NVVNOGTATARNVVNDPVDGYSHASGQVLSFNLDMPGDKKVFYVEFCQRRQIT 298
QY 301 NVATVTVCGGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHVDV 360
Db 299 NIATVTVCGGHKCSANVTTVNPPCVQVNIISGADWSYVCKPVEYSISVSNPGDLVLHVDV 358
```

```
QY 361 IQDTLPSGVTVLEAPGGETCCNKKVVRRIKEMCPGETLOFLKLVKAQVPGPRTNOVAVTSE 420
Db 359 IEDTLSPGIVTVAAGAQAISCNKLKLVMTKELNPGESLQYKVLVRAQTPGQFTNNVVKSC 418
QY 421 SNGCTCTCAEATTHWKGLAATHMCLVLDNDPDCVCGENTVYRICVTVNRGSAEDTNVSLIL 480
Db 419 SDGIGCTCAEATTYWKGAATHMCLVVDTCDPICVCGENTVYRICVTVNRGSAEDTNVSLIL 478
QY 481 KFSKELQPIASSGPTKGTISGNVTFDAPLKLGSKESVEFSVTLKGIAPGDARGEAI 537
Db 479 KFSKELQPIASSGPTKGTITGNVTFDPLRLGSKETVEFSVTLKAYSAGDARGEAILSS 538
QY 541 DLTSPVSDTENTHY 556
Db 539 DLTVPVSDTENTHIY 554

RESULT 6
Q93317 PRELIMINARY: PRT; 547 AA.
AC Q93317;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN.
GN OMPA.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UW-12, AND UW-36;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.B., Tavare S., Dean D.;
RT "Recombination in the ompA Gene but Not the ompB Gene of Chlamydia
RT Contributes to Serovar-Specific Differences in Tissue Tropism, Immune
RT Surveillance, and Persistence of the Organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304330; AAL14100.1; -.
DR EMBL; AF304329; AAL14099.1; -.
SQ SEQUENCE 547 AA; 58626 MW; 817BA3000FEA6A71 CRC64;
```

```
Query Match 73.1%; Score 2157.5; DB 2; Length 547;
Best Local Similarity 71.9%; Pred. No. 3.1e-168;
Matches 402; Conservative 73; Mismatches 69; Indels 15; Gaps 4;

QY 1 MSLKIRRVTVLALTSMAFCASGIEAAVAESLTIKIVASAEKPAVPMTAKKVLRLV 60
Db 1 MNKLIRRAVTIFAVTSVLSFASGLVETSMASLSLTVISLADTKAK--DNTSHKSKAR 58

QY 61 RN--KQPVQKSRGAFCDKEEYPCPEGRCQPVQAEQESCYGLYSVKVNDNCNVEICQS 117
Db 59 KKHSETPVDR-----KEVAPVHESKA--TGPQKQDCSGFRMYTVKVNDRNVEITQA 108

QY 118 VPEYATVGSYPYIEILAIIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRL 177
Db 109 VPEYATVGSYPYIEITATCKRCDVDVITQQLPCEAEFVSSDPATPTADGKLWKIDRL 168

QY 178 GAGDKCKITVWVKLEGGCCFTAAATVACAPELRSYTKGQPAICIKOEGPDCACLRCPVC 237
Db 169 GQGEKSKITVWVKLEGGCCFTAAATVACAPELRSYTKGQPAICIKOEGPDCACLRCPV 228

QY 238 YKIEVNTVGSATARNVTVNDPNPDGYSHASGORVLSFNLGDMRPGDKKVTVEFCQRRG 297
Db 229 YKINVNNOGTARNVTVNDPNPDGYSHASGORVLSFNLGDMRPGDKKVTVEFCQRRG 288

QY 298 QITNATVTCYCGHCKSANVTVVNEPCQVNIISGADWSYVCKPVEYSISVSNPGLVLH 357
Db 289 CATNIATVSYCGHCKNTASVTVINEPCQVNIISGADWSYVCKPVEYSISVSNPGLVLR 348

QY 358 DVIQDTLPSGVTVLEAPGGETCCNKKVVRRIKEMCPGETLOFLKLVKAQVPGPRTNOVAV 417
```

```
Db 349 DVVVEDTLSPGVTVLEAAGAQAISCNKKVWTVVKELNPGESLQYKVLVRAQTPGQFTNNVV 408
QY 418 TSNSCGTCTCAEATTHWKGLAATHMCLVLDNDPDCVCGENTVYRICVTVNRGSAEDTNVS 477
Db 409 KSCDCGCTCTCAEATTYWKGAATHMCLVVDTCDPVCGENTVYRICVTVNRGSAEDTNVS 468
QY 478 LILKFSKELQPIASSGPTKGTISGNVTFDAPLKLGSKESVEFSVTLKGIAPGDARGEAI 537
Db 469 LMLKFSKELQPIASSGPTKGTITGNVTFDPLRLGSKETVEFSVTLKAYSAGDARGEAI 528
QY 538 LSSDTLTSPVSDTENTHY 556
Db 529 LSSDTLTVPVSDTENTHIY 547

RESULT 7
Q9S6B3 PRELIMINARY: PRT; 395 AA.
AC Q9S6B3;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TremBLrel. 13, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR388;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of the
RT outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111201; AAD20337.1; -.
FT NON_TER 1
FT NON_TER 395
FT NON_TER 395
SQ SEQUENCE 395 AA; 42272 MW; 7AD2B76F3F06EA37 CRC64;
```

```
Query Match 72.1%; Score 2127; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 6.5e-166;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAGDKCKITVWVKPLKE 194
Db 1 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAGDKCKITVWVKPLKE 60

QY 195 GCCFTAAATVACAPELRSYTKGQPAICIKOEGPDCACLRCPVCYKTEVNTGSAIARNVT 254
Db 61 GCCFTAAATVACAPELRSYTKGQPAICIKOEGPDCACLRCPVCYKTEVNTGSAIARNVT 120

QY 255 VDNVPDGYSHASGORVLSFNLGDMRPGDKKVTVEFCQRRGQITNATVTCYGGHKCS 314
Db 121 VDNVPDGYSHASGORVLSFNLGDMRPGDKKVTVEFCQRRGQITNATVTCYGGHKCS 180

QY 315 ANVTTVNPEPCQVNIISGADWSYVCKPVEYSISVSNPGLVDLHDVVIQDTLPSGVTVLEA 374
Db 181 ANVTTVNPEPCQVNIISGADWSYVCKPVEYSISVSNPGLVDLHDVVIQDTLPSGVTVLEA 240

QY 375 PGGEICCNKKVVRRIKEMCPGETLOFLKLVKAQVPGPRTNOVAVTSESNCGTCTSCAETTT 434
Db 241 PGGEICCNKKVVRRIKEMCPGETLOFLKLVKAQVPGPRTNOVAVTSESNCGTCTSCAETTT 300

QY 435 HWKGLAATHMCLVLDNDPDCVCGENTVYRICVTVNRGSAEDTNVSLILKFSKELQPIASSGP 494
Db 301 HWKGLAATHMCLVLDNDPDCVCGENTVYRICVTVNRGSAEDTNVSLILKFSKELQPIASSGP 360

QY 495 TKGTISGNVTVFDALPKLGSKESVEFSVTLKGIAP 529
Db 361 TKGTISGNVTVFDALPKLGSKESVEFSVTLKGIAP 395
```



```
RESULT 8
Q93FV8 PRELIMINARY; PRT; 511 AA.
ID Q93FV8;
AC Q93FV8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE 60 KDA CYSTEINE-RICH OMP (FRAGMENT).
GN OMCB.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAR-13;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.L., Tavare S., Dean D.;
RT "Recombination in the ompA Gene but Not the ompB Gene of Chlamydia
RT Contributes to Serovar-Specific Differences in Tissue Tropism, Immune
RT Surveillance, and Persistence of the Organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304332; AAL14102.1; -.
FT NON_TER 511
SQ SEQUENCE 511 AA; 54978 MW; 27555CB67AA9214B CRC64;

Query Match 68.1%; Score 2008.5; DB 2; Length 511;
Best Local Similarity 71.1%; Pred. No. 4.4e-156;
Matches 372; Conservative 71; Mismatches 65; Indels 15; Gaps 4;

QY 1 MSKLIRRVTVLALTSMAFCGIEAANAESLITKIVASAEATKPAVPMTAKKVLVR 60
DB 1 MNKLIRRAVTFATFVASLFAFSGVLETSMAESLSTNVISLADTKAK--DNTSHKSKAR 58
QY 61 RN---KOPVEOKSGAFCDKEFYCEGRQCPVEAQAQESCYGRLYSVKVNDCNVEIQCS 117
DB 59 KNSKETPVDR-----KEVAPVHESKA--TGPQKDSFCFRMYTVKVNDRNVEITQA 108
QY 118 VPEYATVGSPPYIEILGKDKCDVDVITQOLPCEAEFVSDPETPTSDGKLWVKIDRL 177
DB 109 VPEYATVGSPPYIEILATGKRDVDVITQOLPCEAEFVSDPATPTADGKLWVKIDRL 168
QY 178 GAGDKCKITVWVKPKLKEGCCFTATVCACPELRSYTKCGQAICIKQEGPCACLRCPVC 237
DB 169 GQGEKSKITVWVKPKLKEGCCFTATVCACPELRSYTKCGQAICVKQEGPENACLRCPV 228
QY 238 YKIEVNTGSAIARNTVNDVPDGYSHASGORVLSFNLGDMRPGDKKVFVEFCPQRRG 297
DB 229 YKINIVNGTATARNVVENVPDGYAHSSGORVLTFTLGDMPQGEHRTITVEFCPLKRG 288
QY 298 QITNATVTCYGGHKCSANVTYVNEPCVOVNISGADMSYVCKPVEYSISVSNPGDLVLH 357
DB 289 RATNIATVSYCGGHKNTASVTTVINEPCVQVSIAGADMSYVCKPVEYSISVSNPGDLVL 348
QY 358 DVVIQDTLPSGVTVLEAPGEGCCNKVVRKEMCPGETLQFKLVVKAQVPGGRFTNOVAV 417
DB 349 DVVVEDTLSPGVTVLEAGAQISCNKVVYKELNPGESLQYKVLVRAQTPGQFTNNVV 408
QY 418 TSESCNGCTCAEATTHWGLAATHMCLVDNDPICVGTNVRGSAEDTNVSLILKFSKELQ 477
DB 409 KSCSDCGTSCAEATTVWKGAAATVHCVVVDTPVCGVNTVYRICVTNRGSAEDTNVS 468
QY 478 LILKFSKELQIASSGPTKGTISGNTVFPDALPKLGSKESVEF 520
DB 469 LMLKFSKELQPVFSFGPTKGTITGNTVFPDLSPLRGSKESVEF 511

RESULT 9
Q9X4I2 PRELIMINARY; PRT; 395 AA.
ID Q9X4I2;
AC Q9X4I2;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
```

```
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B577;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of the
RT outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111200; AAD20336.1; -.
FT NON_TER 1
FT NON_TER 395
SQ SEQUENCE 395 AA; 42179 MW; FE4D01BBB03B97F5 CRC64;

Query Match 65.0%; Score 1917; DB 2; Length 395;
Best Local Similarity 87.8%; Pred. No. 9.8e-149;
Matches 347; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 135 IGKKDCVDVITQOLPCEAEFVSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKE 194
DB 1 VGKKDCVNVVITQOLPCEVEFVSDPATPTSDSKLIWITDCLGQGEKCKITVWVKPLKE 60
QY 195 GCCCTAATVCACPELRSYTKCGQAICIKQEGPCACLRCPVCYKIEVVNTGSAIARNVT 254
DB 61 GCCCTAATVCACPELRSYTKCGQAICIKQEGPCACLRCPVCYKIEVCNTGSAIARNVV 120
QY 255 VDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVEFCPQRRGOITNATVTCYGGHKCS 314
DB 121 VDNVPDGYTHASGORVLSFNLGDMRPGDKSCFSVEFCPQRRGKITNATVTCYGGHKCS 180
QY 315 ANVTTVNEPCVQVNISGADMSYVCKPVEYSISVSNPGDLVLHDVIODTLPSPGVTVLEA 374
DB 181 ANVTTVNEPCVQVNISGADMSYVCKPVEYTIIVSNLGDGLDLYDVVEDTVPSPGATILEA 240
QY 375 PGGEICCNKVVVRKEMCPGETLQFKLVVKAQVPGGRFTNOVAVTSESCNGCTCAEATTT 434
DB 241 EGAEICCNKVAWCIEKMPGETLQFKLVVKAQSPGKFTNOVVVKTNSDGTCTSCAEATT 300
QY 435 HWKGLAATHMCLVDNDPICVGTNVRGSAEDTNVSLILKFSKELQIASSGP 494
DB 301 HWKGLAATHMCLVDNDPICVGTNVRGSAEDTNVSLILKFSKELQPVSSSGP 360
QY 495 TKGITSGNTVFPDALPKLGSKESVEFVTLKGIAIP 529
DB 361 TKGITGNTVFPDALPKLGSKESVEFVTLKGIAIP 395

RESULT 10
Q9X4I1 PRELIMINARY; PRT; 395 AA.
ID Q9X4I1;
AC Q9X4I1;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCSTRA;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of
RT homologs of the outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111199; AAD20335.1; -.
FT NON_TER 1
FT NON_TER 1
```

```
FT NON_TER 395 395
SQ SEQUENCE 395 AA: 42205 MW: 6A90D7A415A89E CRC64;

Query Match 62.1%; Score 1833; DB 2; Length 395;
Best Local Similarity 83.5%; Pred. No. 7.3e-142;
Matches 330; Conservative 34; Mismatches 31; Indels 0; Gaps 0;

QY 135 IGKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVWVKPLKE 194
Db 1 IGKKRCVNVITQOLPCEAEFVSSDPETPTADGKLMWKIDRLGQGERCKITVWVKPLKE 60

QY 195 GCCTAATVACAPERSYTKGCGPAICIKQEGPCACILRCPCVCYKIEVNTGSAIARNVT 254
Db 61 GCCTAATVACAPERSYTKGCGPAICIKQEGPCACILRCPCVCYKIEVNTGSAIARNV 120

QY 255 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVFTEFCPPRRGQITNVATVTCYCGHKCS 314
Db 121 VDNVPDGFSHASQORVLSFNLGDMHFGESKVLCEFCPPQRGQVTVATVSYCGGHKCS 180

QY 315 ANVTTVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVVIQDTLPSPGVVLEA 374
Db 181 ANVTTVINEPCVNVITISGVDWAYVCKPVEYITVSNPGDLVLNVVVDTLPGSALILEA 240

QY 375 PGGEICCNKVVWRKEMCPGETLQFKLVVKAQVPGRETNQAVTSESNCGTCTCAETTT 434
Db 241 AGAEISCNKAVWCIELCPGETLQFKLVVKAQVPGQITTSVQVNSQDCGACTSCADAT 300

QY 435 HWKGLAATHMCVLDNDPICVGNTRYICVTNRGSAEDTNVSLILKPSKELQPIASSGP 494
Db 301 YWKGLAATHMCVLDNDPICVGNTRYICITNRGSAEDTNVSLILKPSKELQPISSSGP 360

QY 495 TKGTISGNTVFDALPKLGSKESEFVSITLKGIA 529
Db 361 TKGTISGNTVFDALPKLGSKESEFVSITLKGVA 395

RESULT 11
Q46164 ID Q46164 PRELIMINARY; PRT; 183 AA.
AC Q46164;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KC, AND 736;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56925; AAB00572.1; -.
FT NON_TER 1 183
FT NON_TER 183
SQ SEQUENCE 183 AA: 19846 MW: 09B8D084FF4C016E CRC64;

Query Match 32.4%; Score 956; DB 2; Length 183;
Best Local Similarity 98.9%; Pred. No. 1.6e-70;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 ALTNMASCFAAGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVRRNKQPVQKSRG 72
Db 1 ALTNMASCFAAGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVRRNKQPVQKSRG 60

QY 73 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVEICQSVPEVATVGSPIPIEI 132
Db 61 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVEICQSVPEVATVGSPIPIEI 120

Query Match 31.2%; Score 920; DB 2; Length 176;
Best Local Similarity 98.9%; Pred. No. 1.3e-67;
Matches 174; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 ALTNMASCFAAGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVRRNKQPVQKSRG 72
Db 1 ALTNMASCFAAGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVRRNKQPVQKSRG 60

QY 73 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVEICQSVPEVATVGSPIPIEI 132
Db 61 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVEICQSVPEVATVGSPIPIEI 120

QY 133 LAIGKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVW 188
Db 121 LAIGKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVW 176

RESULT 13
Q9AF83 ID Q9AF83 PRELIMINARY; PRT; 182 AA.
AC Q9AF83;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia felis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83356;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP;
RX MEDLINE=21417404; PubMed=11526131;
RA Hartley J.C., Kaye S., Stevenson S., Bennett J., Ridgway G.;
RT "Per detection and molecular identification of Chlamydiaceae species.";
RT species.";
```

```
RL J. Clin. Microbiol. 39:3072-3079(2001).
DR EMBL: AF367407; AAK38115.1; -.
OR InterPro: IPR003506; Chlam_OMP6.
DR PRINTS: PRO1336; CHLAMIDIAOM6.
FT NON_TER 1
FT NON_TER 182
SQ SEQUENCE 182 AA; 19903 MW; 9CEB2D5192EBFICA CRC64;

Query Match      23.2%; Score 685; DB 2; Length 182;
Best Local Similarity 71.4%; Pred. No. 2.3e-48;
Matches 130; Conservative 19; Mismatches 31; Indels 2; Gaps 2;

QY 12 LAITSMASCFASGGIEAAVAESLITKIVASAEKTPAPV-PMTAKKVLVRRNKQPVBEKS 70
DB 1 LAITSMASCFASGGIEAAVAESLITKIVASAEKTPAPV-PMTAKKVLVRRNKQPVBEKS 60
QY 71 RGAFCDKEFYPCGEGRCQ-PVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGSYP 129
DB 61 NNAFCDKEFYPCGEGRCQSSVDTRQESYCGKMSYRVNDDCNVEISQAVPEYATVGSYP 120
QY 130 IEILAIGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVWV 189
DB 121 IEILAIGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVWV 180
QY 190 KP 191
DB 181 KP 182

RESULT 14
QYRB53 PRELIMINARY; PRT; 128 AA.
AC Q9RB53;
AT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN B (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=833558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FROG-MI-L; PubMed=10364623;
RX MEDLINE=99294794; PubMed=10364623;
RA Berger L., Volp K., Mathews S., Speare R., Timms P.;
RT "Chlamydia pneumoniae in a free-ranging giant barred frog (Mixophyes
iteratus) from Australia.";
RL J. Clin. Microbiol. 37:2378-2380(1999).
DR EMBL: AF102831; AAD25999.1; -.
FT NON_TER 1
FT NON_TER 128
SQ SEQUENCE 128 AA; 14112 MW; 0C38639151098A70 CRC64;

Query Match      22.8%; Score 674; DB 2; Length 128;
Best Local Similarity 98.4%; Pred. No. 1.2e-47;
Matches 126; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 38 IVASAEKTPAPVMTAKKVLVRRNKQPVBEKSRGAFCDKEFYPCGEGRCQPVAEQOESC 97
DB 1 IVASAEKTPAPVMTAKKVLVRRNKQPVBEKSRGAFCDKEFYPCGEGRCQPVAEQOESC 60
QY 98 YGRLYSVKYNDDCNVEICQSVPEYATVGSYPPIEILAIGKKDCVDVITQQLPCEAEFVS 157
DB 61 YGRLYSVKYNDDCNVEICQSVPEYATVGSYPPIEILAIGKKDCVDVITQQLPCEAEFVS 120
QY 158 SDPETTPT 165
DB 121 SDPETTPT 128
```

```
RESULT 15
Q46163 PRELIMINARY; PRT; 178 AA.
AC Q46163;
AT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K13, MUT13, MUT133, GC522, AND GC232;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia
pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL: U56924; AAB00571.1; -.
FT NON_TER 1
FT NON_TER 178
SQ SEQUENCE 178 AA; 19490 MW; FF57E90C9D58B691 CRC64;
```

```
Query Match      22.7%; Score 669; DB 2; Length 178;
Best Local Similarity 72.2%; Pred. No. 4.6e-47;
Matches 130; Conservative 21; Mismatches 21; Indels 8; Gaps 2;

QY 13 ALTSMASCFASGGIEAAVAESLITKIVASAEKTPAPVMTAKKVLVRRNKQPVBEK 67
DB 1 ALTSMASCFASGGIEAAVAESLITKIVASAEKTPAPVMTAKKVLVRRNKQPVBEK 57
QY 68 OKSRGAFCDKEFYPCGEGRCQPVBEKSRGAFCDKEFYPCGEGRCQPVBEKSRGAFCDKEFY 127
DB 58 SRSNFCDCDKEFFPCGEGRCQPVBEKSRGAFCDKEFYPCGEGRCQPVBEKSRGAFCDKEFY 117
QY 128 YPIETILAIGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 187
DB 118 YPIETILAIGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 177
```

Search completed: May 25, 2002, 22:23:59  
Job time: 502 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:26:03 ; Search time 31.45 Seconds  
(without alignments)  
3058.352 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 556  
Sequence: 1 MSKLRRVVTVLALTSAMSC.....ILSSDVLTSVPVDIENTHYV 556

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 562222 seqs, 172994929 residues

Word size : 12  
Total number of hits satisfying chosen parameters: 19

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archheap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 395   | 71.0        | 395    | 2 Q9S6B3  | Q9S6B3 chlamydia p |
| 2          | 134   | 24.1        | 183    | 2 Q46164  | Q46164 chlamydia p |
| 3          | 127   | 22.8        | 176    | 2 Q9AG10  | Q9AG10 chlamydia p |
| 4          | 104   | 18.7        | 128    | 2 Q9RB53  | Q9RB53 chlamydia p |
| 5          | 46    | 8.3         | 395    | 2 Q9X412  | Q9X412 chlamydia p |
| 6          | 46    | 8.3         | 534    | 2 Q9ZAW1  | Q9ZAW1 chlamydia p |
| 7          | 46    | 8.3         | 557    | 2 Q9AIS7  | Q9AIS7 chlamydia p |
| 8          | 45    | 8.1         | 558    | 2 P94664  | P94664 chlamydia m |
| 9          | 39    | 7.0         | 554    | 16 Q9PJVO | Q9PJVO chlamydia m |
| 10         | 31    | 5.6         | 534    | 2 Q9ZAW0  | Q9ZAW0 chlamydia p |
| 11         | 26    | 4.7         | 178    | 2 Q46163  | Q46163 chlamydia p |
| 12         | 26    | 4.7         | 178    | 2 Q46165  | Q46165 chlamydia p |
| 13         | 25    | 4.5         | 511    | 2 Q93FV8  | Q93FV8 chlamydia t |
| 14         | 25    | 4.5         | 547    | 2 Q93317  | Q93317 chlamydia t |
| 15         | 24    | 4.3         | 395    | 2 Q9X411  | Q9X411 chlamydia p |
| 16         | 17    | 3.1         | 157    | 2 Q9AF85  | Q9AF85 chlamydia p |

17 3.1 157 2 Q9AF84 chlamydia p  
18 3.1 182 2 Q9AF83 chlamydia p  
19 2.9 173 2 Q9AF82 chlamydia s

#### ALIGNMENTS

##### RESULT 1

Q9S6B3 PRELIMINARY; PRT; 395 AA.  
ID Q9S6B3  
AC Q9S6B3;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).  
GN OMP2.  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AR388;  
RA Kaltenboeck B., Gao D.;  
RT "PCR amplification and sequencing of the partial coding region of the outer membrane protein 2 Omp2 of Chlamydia";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF111201; AAD20337.1; -.  
FT NON\_TER 1  
FT NON\_TER 395  
SQ SEQUENCE 395 AA; 42272 MW; 7AD2B76F3F06EA37 CRC64;

Query Match 71.0%; Score 395; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKE 194  
DB 1 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKE 60

QY 195 GCFTAAATVCACPELRSYTKGQPAICIKQEGDCACLRCPVCYKIEVNTGSAIARNVT 254  
DB 61 GCFTAAATVCACPELRSYTKGQPAICIKQEGDCACLRCPVCYKIEVNTGSAIARNVT 120

QY 255 VDNVPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQITNATVTCGGHKCS 314  
DB 121 VDNVPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQITNATVTCGGHKCS 180

QY 315 ANVTTVNEPCVQVNLISGADWSYCKPVEYSISVSNPGDLVLDVVIQDTLPFGVTVLEA 374  
DB 181 ANVTTVNEPCVQVNLISGADWSYCKPVEYSISVSNPGDLVLDVVIQDTLPFGVTVLEA 240

QY 375 PGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOAVTSESCGTCTSCAETTT 434  
DB 241 PGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOAVTSESCGTCTSCAETTT 300

QY 435 HWKGLAATHMCLVDNDPDCVGNVYRCVTVNRGSAEDTNVSLILKFKSKELOPIASSGP 494  
DB 301 HWKGLAATHMCLVDNDPDCVGNVYRCVTVNRGSAEDTNVSLILKFKSKELOPIASSGP 360

QY 495 TKGTISGNTVVFDPALPKLGSKESEVFSVTLKGIAP 529  
DB 361 TKGTISGNTVVFDPALPKLGSKESEVFSVTLKGIAP 395

##### RESULT 2

Q46164 PRELIMINARY; PRT; 183 AA.  
ID Q46164  
AC Q46164;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KC, AND 736;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia
   pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56925; AAB00572.1; -.
FT NON_TER 1
FT NON_TER 183
SQ SEQUENCE 183 AA; 19846 MW; 09B8D084F4C016E CRC64;

Query Match 24.1%; Score 134; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 6.9e-133;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 109

QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 169

QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

RESULT 3
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 169
QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

Query Match 22.8%; Score 127; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.7e-125;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 109

QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 169

```

```

QY 182 KCKITVW 188
DB 170 KCKITVW 176

RESULT 4
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 165
DB 85 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 128

Query Match 18.7%; Score 104; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.5e-101;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 25 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 84
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 165
DB 85 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 128

RESULT 5
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 25 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 84
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 165
DB 85 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPT 128

Query Match 8.3%; Score 46; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOQESCYGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFFVSSDPETPTSDGKLYWKIDRLGAGD 169

```

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 227  
 Db 48 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 93

RESULT 6  
 Q9ZAW1  
 ID Q9ZAW1 PRELIMINARY; PRT; 534 AA.  
 AC Q9ZAW1;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).  
 OS Chlamydomophila abortus.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
 OX NCBI\_TaxID=83555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S26/3;  
 RA Sheehy N., Markey B., Quinn P.J.;  
 RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes."  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U76760; AAD09597.1; -.  
 DR InterPro; IPR003506; Chlam\_OMP6.  
 DR PRINTS; PR01336; CHLAMIDIOM6.  
 FT NON\_TER 1  
 FT NON\_TER 534  
 SQ SEQUENCE 534 AA; 57137 MW; 27898208041E92BE CRC64;

Query Match 8.3%; Score 46; DB 2; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-39;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 227  
 Db 174 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 219

RESULT 7  
 Q9AIS7  
 ID Q9AIS7 PRELIMINARY; PRT; 557 AA.  
 AC Q9AIS7;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE 60 KDA CYSTEINE-RICH MEMBRANE COMPLEX PROTEIN.  
 GN CMCB.  
 OS Chlamydomophila abortus.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
 OX NCBI\_TaxID=83555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EBA;  
 RX MEDLINE=21078680; PubMed=11211261;  
 RA Bush R.M., Everett K.D.;  
 RT "Molecular evolution of the Chlamydiaceae."  
 RL Int. J. Syst. Evol. Microbiol. 51:203-220(2001).  
 DR EMBL; AF240773; AAG60550.1; -.  
 DR InterPro; IPR003506; Chlam\_OMP6.  
 DR PRINTS; PR01336; CHLAMIDIOM6.  
 SQ SEQUENCE 557 AA; 59761 MW; 817A20B071AD9609 CRC64;

Query Match 8.3%; Score 46; DB 2; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-39;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 227  
 Db 183 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 228

RESULT 8  
 P94664  
 ID P94664 PRELIMINARY; PRT; 558 AA.  
 AC P94664;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE OUTER MEMBRANE PROTEIN 2.  
 GN OMP2.  
 OS Chlamydomophila caviae.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
 OX NCBI\_TaxID=83557;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS (GPIC);  
 RX MEDLINE=97075924; PubMed=8918247;  
 RA Hsia R.C., Bavoil P.M.;  
 RT "Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: structural and functional implications."  
 RL Gene 176:155-162(1996).  
 DR EMBL; U41759; AAB41143.1; -.  
 DR InterPro; IPR003506; Chlam\_OMP6.  
 DR PRINTS; PR01336; CHLAMIDIOM6.  
 SQ SEQUENCE 558 AA; 60213 MW; 452E2BD79078935C CRC64;

Query Match 8.1%; Score 45; DB 2; Length 558;  
 Best Local Similarity 100.0%; Pred. No. 2e-38;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 227  
 Db 185 KCKITVWVKPLKEGCGCFTAAATVCACPRLSYTKGQPAICIKQEGP 229

RESULT 9  
 Q9PJVO  
 ID Q9PJVO PRELIMINARY; PRT; 554 AA.  
 AC Q9PJVO;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE 60 KDA OUTER MEMBRANE PROTEIN.  
 GN TC0727.  
 OS Chlamydia muridarum.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MOPN / NIGG;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AK39.";  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 DR EMBL; AE002341; AAF39537.1; -.  
 DR TIGR; TC0727; -.  
 KW Complete proteome.  
 SQ SEQUENCE 554 AA; 59784 MW; 1CF81B2471C27FD4 CRC64;

Query Match 7.0%; Score 39; DB 16; Length 554;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-32;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 DPICVGVTVYRICVTVNRGSAEDTNVSLILKFSKELOPI 489  
 Db 449 DPICVGVTVYRICVTVNRGSAEDTNVSLILKFSKELOPI 487

```

RESULT 10
Q9ZAW0 ID Q9ZAW0 PRELIMINARY; PRT; 534 AA.
AC Q9ZAW0:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE CYS-TEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR73;
RA Sheehy N., Markey B., Quinn P.J.;
RT "Sequence analysis of C. psittaci and C. pecorum 60 kda genes.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DDJ databases.
DR EMBL; U76761; AAD09598.1; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
FT NON_TER 1 534
FT NON_TER 534 534
SQ SEQUENCE 534 AA; 57298 MW; 3929DBD6A8534258 CRC64;

Query Match 5.6%; Score 31; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 506 FDALPKLGSKEVSFVTLKGIAPGDARGE 536
Db 498 FDALPKLGSKEVSFVTLKGIAPGDARGE 528

RESULT 11
Q46163 ID Q46163 PRELIMINARY; PRT; 178 AA.
AC Q46163:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE OMPB (FRAGMENT).
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K13, MUT13, MUT133, GC522, AND GC232;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56924; AAB00571.1; -.
FT NON_TER 1 178
FT NON_TER 178 178
SQ SEQUENCE 178 AA; 19490 MW; FF57E90C9D58B691 CRC64;

Query Match 4.7%; Score 26; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 8.8e-19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 QSVPEYATVGSYPPIEILAIKDKCV 141
Db 106 QSVPEYATVGSYPPIEILAIKDKCV 131

RESULT 12
Q46165 ID Q46165 PRELIMINARY; PRT; 178 AA.

```

```

AC Q46165;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR268;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56927; AAB00574.1; -.
FT NON_TER 1 178
FT NON_TER 178 178
SQ SEQUENCE 178 AA; 19506 MW; E296350C9D58B691 CRC64;

Query Match 4.7%; Score 26; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 8.8e-19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 QSVPEYATVGSYPPIEILAIKDKCV 141
Db 106 QSVPEYATVGSYPPIEILAIKDKCV 131

RESULT 13
Q93FV8 ID Q93FV8 PRELIMINARY; PRT; 511 AA.
AC Q93FV8:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 60 KDA CYSTEINE-RICH OMP (FRAGMENT).
GN OMCB.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAR-13;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.L., Tavares S., Dean D.;
RT "Recombination in the ompA Gene but Not the ompB Gene of Chlamydia RT Surveillance, and Persistence of the Organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304332; AAL14102.1; -.
FT NON_TER 511 511
SQ SEQUENCE 511 AA; 54978 MW; 27555CB67AA9214B CRC64;

Query Match 4.5%; Score 25; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 KITVWVKPLKEGCCFTAAATVCACPE 208
Db 175 KITVWVKPLKEGCCFTAAATVCACPE 199

RESULT 14
Q93317 ID Q93317 PRELIMINARY; PRT; 547 AA.
AC Q93317:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```



Search completed: May 25, 2002, 22:29:44  
Job time: 221 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 22:24:57 ; Search time 118.77 Seconds  
(without alignments)  
1647.733 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 556  
Sequence: 1 MSKLIRRVTVLALTSMA...ILSSDLTSPVSDTENTHYV 556

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 12

Total number of hits satisfying chosen parameters: 67

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 556   | 100.0       | 556    | 15    | US-09-198-452A-596 |
| 2          | 556   | 100.0       | 556    | 15    | US-09-523-647-2    |
| 3          | 556   | 100.0       | 556    | 22    | US-09-841-132-398  |
| 4          | 556   | 100.0       | 559    | 18    | US-09-438-185-559  |
| 5          | 556   | 100.0       | 559    | 18    | US-09-438-185A-559 |
| 6          | 25    | 4.5         | 25     | 24    | US-10-020-269-38   |
| 7          | 25    | 4.5         | 553    | 22    | US-09-841-132-441  |

|                       |    |                     |                                  |               |               |
|-----------------------|----|---------------------|----------------------------------|---------------|---------------|
| Query Match           |    | 100.0%;             | Score 556;                       | DB 15;        | Length 556;   |
| Best Local Similarity |    | 100.0%;             | Pred. No. 0;                     |               |               |
| Matches 556;          |    | Conservative 0;     | Mismatches 0;                    | Indels 0;     | Gaps 0;       |
| Qy                    | 1  | MSKLIRRVTVLALTSMA   | SCFASGGIEAAVAESLTKIVASAETKPAPVPM | AKKRVLR       | 60            |
| Db                    | 1  | MSKLIRRVTVLALTSMA   | SCFASGGIEAAVAESLTKIVASAETKPAPVPM | AKKRVLR       | 60            |
| Qy                    | 61 | RNKQPEOKSRGAFCDKEFY | PCPEGRCQVPEAQQESC                | YGRLYSVKVNDCN | VEICQSVPE 120 |
| Db                    | 61 | RNKQPEOKSRGAFCDKEFY | PCPEGRCQVPEAQQESC                | YGRLYSVKVNDCN | VEICQSVPE 120 |

ALIGNMENTS

RESULT 1  
US-09-198-452A-596  
; Sequence 596, Application US/09198452A  
; GENERAL INFORMATION:  
; APPLICANT: Griflais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198.452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 596  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-596

|    |    |     |     |    |                     |
|----|----|-----|-----|----|---------------------|
| 8  | 25 | 4.5 | 557 | 16 | US-09-201-228A-1036 |
| 9  | 20 | 3.6 | 20  | 24 | US-10-020-269-41    |
| 10 | 19 | 3.4 | 19  | 24 | US-10-020-269-39    |
| 11 | 16 | 2.9 | 16  | 18 | US-09-454-684-253   |
| 12 | 16 | 2.9 | 16  | 18 | US-09-454-684A-253  |
| 13 | 16 | 2.9 | 16  | 18 | US-09-461-705-7     |
| 14 | 16 | 2.9 | 16  | 19 | US-09-556-877-253   |
| 15 | 16 | 2.9 | 16  | 19 | US-09-598-419-253   |
| 16 | 16 | 2.9 | 16  | 20 | US-09-620-412A-253  |
| 17 | 16 | 2.9 | 16  | 22 | US-09-841-132-253   |
| 18 | 16 | 2.9 | 16  | 24 | US-10-020-269-43    |
| 19 | 16 | 2.9 | 20  | 18 | US-09-454-684-250   |
| 20 | 16 | 2.9 | 20  | 18 | US-09-454-684A-250  |
| 21 | 16 | 2.9 | 20  | 19 | US-09-556-877-250   |
| 22 | 16 | 2.9 | 20  | 19 | US-09-598-419-250   |
| 23 | 16 | 2.9 | 20  | 20 | US-09-620-412A-250  |
| 24 | 16 | 2.9 | 20  | 22 | US-09-841-132-250   |
| 25 | 13 | 2.3 | 19  | 24 | US-10-020-269-40    |
| 26 | 12 | 2.2 | 12  | 18 | US-09-454-684-252   |
| 27 | 12 | 2.2 | 12  | 18 | US-09-454-684A-252  |
| 28 | 12 | 2.2 | 12  | 19 | US-09-556-877-252   |
| 29 | 12 | 2.2 | 12  | 19 | US-09-598-419-252   |
| 30 | 12 | 2.2 | 12  | 20 | US-09-620-412A-252  |
| 31 | 12 | 2.2 | 12  | 22 | US-09-841-132-252   |
| 32 | 12 | 2.2 | 16  | 18 | US-09-454-684-251   |
| 33 | 12 | 2.2 | 16  | 18 | US-09-454-684A-251  |
| 34 | 12 | 2.2 | 16  | 19 | US-09-556-877-251   |
| 35 | 12 | 2.2 | 16  | 19 | US-09-598-419-251   |
| 36 | 12 | 2.2 | 16  | 20 | US-09-620-412A-251  |
| 37 | 12 | 2.2 | 16  | 22 | US-09-841-132-251   |
| 38 | 12 | 2.2 | 20  | 18 | US-09-454-684-241   |
| 39 | 12 | 2.2 | 20  | 18 | US-09-454-684-242   |
| 40 | 12 | 2.2 | 20  | 18 | US-09-454-684-246   |
| 41 | 12 | 2.2 | 20  | 18 | US-09-454-684A-246  |
| 42 | 12 | 2.2 | 20  | 18 | US-09-454-684-249   |
| 43 | 12 | 2.2 | 20  | 18 | US-09-454-684A-241  |
| 44 | 12 | 2.2 | 20  | 18 | US-09-454-684A-242  |
| 45 | 12 | 2.2 | 20  | 18 | US-09-454-684A-246  |



Db 241 EVVNTGSAIARNVTVDNPPVPGDYGSHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
QY 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360  
Db 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360  
QY 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420  
Db 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420  
QY 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVYCNRGSADETNVSLL 480  
Db 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVYCNRGSADETNVSLL 480  
QY 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540  
Db 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540  
QY 541 DTLTSPVSDTENTHVV 556  
Db 541 DTLTSPVSDTENTHVV 556

## RESULT 4

US-09-438-185-559  
; Sequence 559, Application US/09438185  
; GENERAL INFORMATION:  
; APPLICANT: Stephens, Richard  
; APPLICANT: Mitchell, Wayne  
; APPLICANT: Kalman, Sue  
; APPLICANT: Davis, Ronald  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
; FILE REFERENCE: 018941-000411US  
; CURRENT APPLICATION NUMBER: US/09/438,185  
; PRIOR FILING DATE: 1999-11-11  
; PRIOR APPLICATION NUMBER: US 60/108,279  
; PRIOR FILING DATE: 1998-11-12  
; PRIOR APPLICATION NUMBER: US 60/128,606  
; PRIOR FILING DATE: 1999-04-08  
; NUMBER OF SEQ ID NOS: 1074  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 559  
; LENGTH: 559  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-438-185-559

Query Match 100.0%; Score 556; DB 18; Length 559;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKLIRRVVTVLALTSMASCFASGGTEAAVAESLITKIVASAETKPAVPMTAKKVLRLV 60  
Db 4 MSKLIRRVVTVLALTSMASCFASGGTEAAVAESLITKIVASAETKPAVPMTAKKVLRLV 63  
QY 61 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
Db 64 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDDCNVEICQSVPE 123  
QY 121 YATVGSPIYPIETILAIKDKCDVDVITQQLPCEAEFYSSDPETPTSDGKLWVKIDRLGAG 180  
Db 124 YATVGSPIYPIETILAIKDKCDVDVITQQLPCEAEFYSSDPETPTSDGKLWVKIDRLGAG 183  
QY 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGOPAICIKQEGPDCACLRCPVCYKI 240  
Db 184 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGOPAICIKQEGPDCACLRCPVCYKI 243  
QY 241 EVVNTGSAIARNVTVDNPPVPGDYGSHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 244 EVVNTGSAIARNVTVDNPPVPGDYGSHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

QY 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360  
Db 304 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 363  
QY 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420  
Db 364 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 423  
QY 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVYCNRGSADETNVSLL 480  
Db 424 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVYCNRGSADETNVSLL 483  
QY 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540  
Db 484 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 543  
QY 541 DTLTSPVSDTENTHVV 556  
Db 544 DTLTSPVSDTENTHVV 559

## RESULT 5

US-09-438-185A-559  
; Sequence 559, Application US/09438185A  
; GENERAL INFORMATION:  
; APPLICANT: Stephens, Richard  
; APPLICANT: Mitchell, Wayne  
; APPLICANT: Kalman, Sue  
; APPLICANT: Davis, Ronald  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
; FILE REFERENCE: 018941-000411US  
; CURRENT APPLICATION NUMBER: US/09/438,185A  
; PRIOR FILING DATE: 2002-03-13  
; PRIOR APPLICATION NUMBER: US 60/108,279  
; PRIOR FILING DATE: 1998-11-12  
; PRIOR APPLICATION NUMBER: US 60/128,606  
; PRIOR FILING DATE: 1999-04-08  
; NUMBER OF SEQ ID NOS: 1074  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 559  
; LENGTH: 559  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; OTHER INFORMATION: CPh0557  
US-09-438-185A-559

Query Match 100.0%; Score 556; DB 18; Length 559;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKLIRRVVTVLALTSMASCFASGGTEAAVAESLITKIVASAETKPAVPMTAKKVLRLV 60  
Db 4 MSKLIRRVVTVLALTSMASCFASGGTEAAVAESLITKIVASAETKPAVPMTAKKVLRLV 63  
QY 61 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDDCNVEICQSVPE 120  
Db 64 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDDCNVEICQSVPE 123  
QY 121 YATVGSPIYPIETILAIKDKCDVDVITQQLPCEAEFYSSDPETPTSDGKLWVKIDRLGAG 180  
Db 124 YATVGSPIYPIETILAIKDKCDVDVITQQLPCEAEFYSSDPETPTSDGKLWVKIDRLGAG 183  
QY 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGOPAICIKQEGPDCACLRCPVCYKI 240  
Db 184 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGOPAICIKQEGPDCACLRCPVCYKI 243  
QY 241 EVVNTGSAIARNVTVDNPPVPGDYGSHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300  
Db 244 EVVNTGSAIARNVTVDNPPVPGDYGSHASQORVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

```
QY 301 NVATVTCGGHKCSANVTTVNPECVQVNIISGADWSYCKPVEYSISVSNPGLVLDHV 360
|||||
Db 304 NVATVTCGGHKCSANVTTVNPECVQVNIISGADWSYCKPVEYSISVSNPGLVLDHV 363
QY 361 IQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVVRAQVPGRTNQAVTSE 420
|||||
Db 364 IQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVVRAQVPGRTNQAVTSE 423
QY 421 SNGCTCTSCAETTHWKGLAATHMCLVLDNDPICVGVNTVYRICVTNRGSAEDTNVSLIL 480
|||||
Db 424 SNGCTCTSCAETTHWKGLAATHMCLVLDNDPICVGVNTVYRICVTNRGSAEDTNVSLIL 483
QY 481 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540
|||||
Db 484 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFVTLKGIAPGDARGEAILSS 543
QY 541 DTLTSPVSDTENTHVV 556
|||||
Db 544 DTLTSPVSDTENTHVV 559
|||||
RESULT 6
US-10-020-269-38
; Sequence 38, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-020-269-38
```

```
Query Match 4.5%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RRNKQPVQKSRGAFCDKEFYPCEE 84
|||||
Db 1 RRNKQPVQKSRGAFCDKEFYPCEE 25
```

```
RESULT 7
US-09-841-132-441
; Sequence 441, Application US/09841132
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 441
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-441
```

```
Query Match 4.5%; Score 25; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAATVCAAPE 208
|||||
Db 181 KITVWVKPLKEGCGCFTAATVCAAPE 205
```

```
RESULT 8
US-09-201-228A-1036
; Sequence 1036, Application US/09201228A
; GENERAL INFORMATION:
; APPLICANT: Griffais, Remy
; APPLICANT: Hoiseth, Susan K.
; APPLICANT: Zagursky, Robert John
; APPLICANT: Metcalf, Benjamin J.
; APPLICANT: Peek, Joel A.
; APPLICANT: Sankaran, Banumathi
; APPLICANT: Fletcher, Leah Diane
; TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE
; TITLE OF INVENTION: AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN
; TITLE OF INVENTION: PARTICULAR FOR THE DIAGNOSIS, PREVENTION AND TREATMENT OF
; FILE REFERENCE: 9710-0004-999
; CURRENT APPLICATION NUMBER: US/09/201,228A
; CURRENT FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/107,077
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: FR 97-16034
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: FR 97-15041
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 5981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1036
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-201-228A-1036
```

```
Query Match 4.5%; Score 25; DB 16; Length 557;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAATVCAAPE 208
|||||
Db 181 KITVWVKPLKEGCGCFTAATVCAAPE 205
```

```
RESULT 9
US-10-020-269-41
; Sequence 41, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 20
; TYPE: PRT
```

```

; ORGANISM: Chlamydia pneumoniae
US-10-020-269-41

Query Match          3.6%; Score 20; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.7e-13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 418 TSESNGTCTCTCAETTHWK 437
| | | | | | | | | | | | | | | |
Db 1 TSESNGTCTCTCAETTHWK 20

RESULT 10
US-10-020-269-39
; Sequence 39, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; CURRENT FILING DATE: 2001-12-14
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-020-269-39

Query Match          3.4%; Score 19; DB 24; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 DMRPGDKKVFVTFPCQRR 296
| | | | | | | | | | | | | | | |
Db 1 DMRPGDKKVFVTFPCQRR 19

RESULT 11
US-09-454-684-253
; Sequence 253, Application US/09454684
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C4
; CURRENT APPLICATION NUMBER: US/09/454,684
; CURRENT FILING DATE: 1999-12-03
; PRIOR FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-454-684-253

Query Match          2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GIEAAVAESLITKIVA 40
| | | | | | | | | | | | | | | |
Db 1 GIEAAVAESLITKIVA 16

US-09-454-684-253
; Sequence 253, Application US/09454684
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C4
; CURRENT APPLICATION NUMBER: US/09/454,684
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 253
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-454-684-253

Query Match          2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GDKCKITVWVKPLKEG 195
| | | | | | | | | | | | | | | |
Db 1 GDKCKITVWVKPLKEG 16

RESULT 12
US-09-454-684A-253
; Sequence 253, Application US/09454684A
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C4
; CURRENT APPLICATION NUMBER: US/09/454,684A
; CURRENT FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 253
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-454-684A-253

Query Match          2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GDKCKITVWVKPLKEG 195
| | | | | | | | | | | | | | | |
Db 1 GDKCKITVWVKPLKEG 16

RESULT 13
US-09-461-705-7
; Sequence 7, Application US/09461705
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: Novel Peptides Capable of Modulating Inflammatory Heart
; TITLE OF INVENTION: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/461,705
; CURRENT FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 09/133,774
; PRIOR FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-461-705-7

Query Match          2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GIEAAVAESLITKIVA 40
| | | | | | | | | | | | | | | |
Db 1 GIEAAVAESLITKIVA 16
```

Search completed: May 25, 2002, 22:28:39  
Job time: 222 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:25:37 ; Search time 13.95 Seconds  
(without alignments)  
1128.285 Million cell updates/sec

Title: US-09-523-647-2  
Perfect score: 556  
Sequence: 1 MSKLIRRVTVLALTMASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 116914 seqs, 28308587 residues

Word size : 12

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Pending\_Patents\_AA\_New.\*
- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Prei. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | DB Length | ID    | Description |
|------------|-------|-------------|-----------|-------|-------------|
| -----      | ----- | -----       | -----     | ----- | -----       |

No matches found

Search completed: May 25, 2002, 22:29:01  
Job time: 204 sec

